



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 152272

TO: Nancy Vogel
Location: rem/2a70/2c70
Art Unit: 1636
Wednesday, May 11, 2005

Case Serial Number: 09/896888

From: Edward Hart
Location: Biotech-Chem Library
REM-1A55
Phone: 571-272-2512

edward.hart@uspto.gov

Search Notes

Examiner Vogel,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart



This Page Blank (uspto)

STIC-Biotech/ChemLib

152272

mej

From: Vogel, Nancy
Sent: Tuesday, May 03, 2005 7:14 AM
To: STIC-Biotech/ChemLib
Subject: sequence search for 09/896,888

please do sequence search of SEQ ID NO:1, nucleotides 351-527 of SEQ ID NO:1, and SEQ ID NO:10,
including interference files,
of 09/896,888
and return results to me on paper asap.
Thanks

Nancy Vogel
Patent Examiner
Art Unit 1636
Office: Remson 2A70
Mail Box: Remson 2C70
(571) 272-0780

RECEIVED
MAY -3 2005
STIC/BIOTECH DIVISION
(STIC)

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: 5/9/05
Date Completed: 5/11/05
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: 3 AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: OSA
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)



STIC SEARCH RESULTS

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library-Remsen Bldg.



This Page Blank (uspto)

This Page Blank (uspto)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	50	100.0	50	6	BD070868	Insect ex
2	36.8	73.6	99591	2	AC013392	BD070868
3	36.8	73.6	118229	2	AC013315	AC013392 Homo sapi
4	36.8	73.6	11967	2	AC013368	AC013315 Homo sapi
5	36.8	73.6	260636	2	AC015852	AC013368 Homo sapi
6	36.2	72.4	73282	2	AC010800	AC015852 Homo sapi
7	36.2	72.4	217141	2	AC073998	AC010800 Homo sapi
8	35.8	71.6	604	11	G39061	AC073998 Homo sapi
9	35.8	71.6	745	6	BD269301	G39061 Z11905 Zeb
10	35.8	71.6	826	11	G40613	BD269301 33 human
11	35.8	71.6	979	11	G40162	G40613 Z7956 Zeb
12	35.8	71.6	76295	2	AC023542	G40162 Z20177 Zeb
13	35.6	71.2	586	6	CQ406071	AC023542 Homo sapi
14	35.6	71.2	75002	2	AC023384	CQ406071 Sequence
15	35.6	71.2	95127	2	AC016798	AC023384 Homo sapi
16	35.6	71.2	115468	2	AC011820	AC016798 Homo sapi
17	35.6	71.2	170427	2	AC006095	AC011820 Homo sapi
18	35.2	70.4	114	6	AR092751	AC006095 Homo sapi
19	35.2	70.4	114	6	AR359133	AR092751 Sequence

RESULT 2

AC013392

LOCUS

DEFINITION

AC013392 Homo sapiens chromosome 2 clone RP11-429N24 map 2, LOW-PASS

SEQUENCE

AC013392

VERSION

AC013392.3 GI:9123920

KEYWORDS

HTG; HTGS_PHASE0.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,

Baldwin, J., Barna, L., Becker, R., Boguski, L., Bouck, G., Breen, J.,

Brown, A., Castle, A., Collins, S., Collins, S., Collins, S., Collins, S.,

Cooke, P., DeArnell, K., Dewar, K., Domino, M., Donnell, L., Doyle, M.,

Ferrel, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,

Galagan, J., Gardy, S., Grant, G., Hagg, B., Hagg, B., Hagg, B.,

Howland, J. C., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J.,

Lehoczky, J., Lieu, C., Locke, R., Macdonald, P., Marquis, N.,

McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,

Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,

Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,

Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,

Testa, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,

Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission

TITLE

Submitted (09-NOV-1999) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Jul 13, 2000 this sequence version replaced gi:6425709.

All repeats were identified using RepeatMasker:

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Smit, A.F.A. & Green, P. (1996-1997)

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L3805

Center clone name: 429_N_24

* NOTE: This record contains 111 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 827: contig of 827 bp in length

828 927: gap of 100 bp

928 1726: contig of 799 bp in length

1727 1826: gap of 100 bp

1827 2635: contig of 809 bp in length

2636 2736: gap of 100 bp

2736 3547: contig of 812 bp in length

3547 3647: gap of 100 bp

3647 4424: contig of 777 bp in length

4424 5303: gap of 100 bp

5303 5403: contig of 779 bp in length

5403 6171: gap of 100 bp

6171 6171: contig of 768 bp in length

6172 6271: gap of 100 bp
 6272 7071: contig of 800 bp in length
 7072 7171: gap of 100 bp
 7172 7979: contig of 808 bp in length
 7980 8079: gap of 100 bp
 8080 8881: contig of 802 bp in length
 8882 8981: gap of 100 bp
 8982 9732: contig of 751 bp in length
 9733 9832: gap of 100 bp
 9833 10559: contig of 727 bp in length
 10560 11463: contig of 804 bp in length
 11464 12340: contig of 777 bp in length
 12341 12440: gap of 100 bp
 12441 13237: contig of 797 bp in length
 13238 13337: gap of 100 bp
 13339 14134: contig of 797 bp in length
 14135 14234: gap of 100 bp
 14235 15002: contig of 768 bp in length
 15003 15102: gap of 100 bp
 15104 15872: contig of 770 bp in length
 15873 15972: gap of 100 bp
 15973 16731: contig of 759 bp in length
 16732 16831: gap of 100 bp
 16832 17605: contig of 774 bp in length
 17606 17705: gap of 100 bp
 17707 18499: contig of 794 bp in length
 18500 18599: gap of 100 bp
 18600 19401: contig of 802 bp in length
 19402 19501: gap of 100 bp
 19502 20325: contig of 824 bp in length
 20326 20425: gap of 100 bp
 20426 21189: contig of 764 bp in length
 21190 22048: contig of 759 bp in length
 22049 22148: gap of 100 bp
 22149 22907: contig of 759 bp in length
 22908 23007: gap of 100 bp
 23009 23775: contig of 768 bp in length
 23776 23875: gap of 100 bp
 23876 24684: contig of 809 bp in length
 24685 24784: gap of 100 bp
 24785 25560: contig of 776 bp in length
 25561 25660: gap of 100 bp
 25661 26446: contig of 786 bp in length
 26447 26547: gap of 100 bp
 26548 27322: contig of 776 bp in length
 27323 27422: gap of 100 bp
 27423 28179: contig of 757 bp in length
 28180 28279: gap of 100 bp
 28280 29044: contig of 765 bp in length
 29045 29144: gap of 100 bp
 29145 29942: contig of 798 bp in length
 29943 30042: gap of 100 bp
 30043 30811: contig of 769 bp in length
 30812 30911: gap of 100 bp
 30912 31663: contig of 752 bp in length
 31664 32540: contig of 777 bp in length
 32541 32640: gap of 100 bp
 32641 33471: contig of 831 bp in length
 33472 33571: gap of 100 bp
 33572 34358: contig of 787 bp in length
 34359 34458: gap of 100 bp
 34459 35239: contig of 781 bp in length
 35240 35339: gap of 100 bp
 35340 36111: contig of 772 bp in length
 36112 36211: gap of 100 bp
 36212 36966: contig of 755 bp in length
 36967 37066: gap of 100 bp
 37067 37813: contig of 747 bp in length
 37814 37913: gap of 100 bp

```

* 37914 38694: contig of 781 bp in length
* 38795 38794: gap of 100 bp
* 38795 39559: contig of 765 bp in length
* 39560 39559: gap of 100 bp
* 40413 40413: contig of 754 bp in length
* 40414 40513: gap of 100 bp
* 40514 41335: contig of 822 bp in length
* 41336 41435: gap of 100 bp
* 41436 42174: contig of 739 bp in length
* 42175 42274: gap of 100 bp
* 42275 43080: contig of 806 bp in length
* 43081 43180: gap of 100 bp
* 43181 43971: contig of 791 bp in length
* 43972 44071: gap of 100 bp
* 44072 44876: contig of 805 bp in length
* 44877 44976: gap of 100 bp
* 44977 45791: contig of 815 bp in length
* 45792 45891: gap of 100 bp
* 45892 46720: contig of 829 bp in length
* 46721 46820: gap of 100 bp
* 46821 47558: contig of 738 bp in length
* 47559 47658: gap of 100 bp
* 47659 48475: contig of 817 bp in length
* 48476 48575: gap of 100 bp
* 48576 49384: contig of 809 bp in length
* 49385 49484: gap of 100 bp
* 49485 50346: contig of 862 bp in length
* 50347 50446: gap of 100 bp
* 50447 51295: contig of 849 bp in length
* 51296 51395: gap of 100 bp
* 51396 52202: contig of 807 bp in length
* 52203 52302: gap of 100 bp
* 52303 53120: contig of 818 bp in length
* 53121 53220: gap of 100 bp
* 53221 54050: contig of 830 bp in length
* 54051 54150: gap of 100 bp
* 54151 54967: contig of 817 bp in length
* 54968 55067: gap of 100 bp
* 55068 55868: contig of 801 bp in length
* 55869 55968: gap of 100 bp
* 55969 56805: contig of 837 bp in length
* 56806 56905: gap of 100 bp
* 56906 57747: contig of 842 bp in length
* 57748 57847: gap of 100 bp
* 57848 58650: contig of 803 bp in length
* 58651 58750: gap of 100 bp
* 58751 59553: contig of 803 bp in length
* 59554 59653: gap of 100 bp
* 59654 60386: contig of 733 bp in length
* 60387 60486: gap of 100 bp
* 60487 61311: contig of 825 bp in length
* 61312 61411: gap of 100 bp
* 61412 62225: contig of 814 bp in length
* 62226 62325: gap of 100 bp
* 62326 63128: contig of 803 bp in length
* 63129 63228: gap of 100 bp
* 63229 64055: contig of 827 bp in length
* 64056 64155: gap of 100 bp

```

```

Query Match 73.6%; Score 36.8; DB 2; Length 99591;
Best Local Similarity 85.4%; Pred. NO. 0.0031;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy 3 TTAGCTTATAGCATGCTCCCGCTTCCAGTCGGGAACCTGTCG 50
|||||
Db 3876 TTAATTGCTGCGCTGACTCCCGCTTCCAGTCGGGAACCTGTCG 3923

```

RESULT 3

```

AC013315 AC013315 118229 bp DNA linear HTG 24-AUG-2002
LOCUS Homo sapiens chromosome 22 clone RP11-348B6 map 22, LOW-PASS
DEFINITION SEQUENCE SAMPLING.

```

ACCESSION VERSION KEYWORDS SOURCE ORGANISM

AC013315
AC013315.3 GI:9123904
HTG; HTGS PHASE0.
Homo sapiens (human)
Homo sapiens

REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS

1 (bases 1 to 118229)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 22, clone RP11-348B6
Unpublished
2 (bases 1 to 118229)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckerly, R., Boguslavskiy, L., Boukhalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collamore, A.,
Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C.,
Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehocsky, J.,
Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A.,
McKernan, K., McLaughlin, J., Meldrim, J., Morrow, J., Naylor, J., Norman, C. H.,
O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A.,
Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D.,
Ye, W. J., Zimmer, A. and Zody, M.

TITLE JOURNAL REFERENCE AUTHORS

Direct Submission
Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 118229)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F.,
Boguslavskiy, L., Boukhalter, B., Brown, A., Burkett, G., Campopiano, A.,
Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collamore, A.,
Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Gage, D., Dodge, S.,
Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J.,
Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G.,
Hagos, B., Heaford, A., Horton, L., Howland, J. C., Iliiev, I., Johnson, R.,
Jones, C., Kann, L., Karatas, A., Klein, J., LaRoque, K., Lamazares, R.,
Landers, T., Lehocsky, J., Levine, R., Lieu, C., Liu, G., Locke, K.,
Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A.,
McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neil, D., Olivari, T. M., Oliver, J., Peterson, K., Pierre, N., Pisani, C.,
Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A.,
Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A.
and Zody, M.

TITLE JOURNAL COMMENT

Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6425714.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIGB
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3764
Center clone name: 348_B_6

* NOTE: This record contains 124 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows

* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 765: contig of 765 bp in length
* 766: gap of 100 bp
* 865: contig of 811 bp in length
* 1676: gap of 100 bp
* 1677: contig of 801 bp in length
* 2577: contig of 801 bp in length
* 2578: gap of 100 bp
* 2678: contig of 842 bp in length
* 3519: gap of 100 bp
* 3520: contig of 822 bp in length
* 3620: gap of 100 bp
* 4442: contig of 822 bp in length
* 4542: gap of 100 bp
* 5306: contig of 764 bp in length
* 5406: gap of 100 bp
* 6261: contig of 856 bp in length
* 6361: gap of 100 bp
* 6362: contig of 779 bp in length
* 7140: gap of 100 bp
* 7240: contig of 801 bp in length
* 8021: contig of 781 bp in length
* 8121: gap of 100 bp
* 8122: contig of 804 bp in length
* 8926: gap of 100 bp
* 9026: contig of 815 bp in length
* 9841: gap of 100 bp
* 9941: contig of 777 bp in length
* 10718: gap of 100 bp
* 10818: contig of 800 bp in length
* 11618: gap of 100 bp
* 11718: contig of 809 bp in length
* 12527: contig of 100 bp
* 12627: gap of 100 bp
* 13507: contig of 880 bp in length
* 13606: gap of 100 bp
* 14393: contig of 787 bp in length
* 14494: gap of 100 bp
* 15277: contig of 783 bp in length
* 15377: gap of 100 bp
* 16163: contig of 786 bp in length
* 16263: gap of 100 bp
* 17046: contig of 783 bp in length
* 17146: gap of 100 bp
* 17146: contig of 821 bp in length
* 17967: gap of 100 bp
* 18067: contig of 800 bp in length
* 18867: gap of 100 bp
* 18967: contig of 887 bp in length
* 19854: gap of 100 bp
* 19954: contig of 812 bp in length
* 20766: gap of 100 bp
* 20866: contig of 778 bp in length
* 21644: gap of 100 bp
* 21744: contig of 837 bp in length
* 22581: gap of 100 bp
* 22681: contig of 820 bp in length
* 23501: gap of 100 bp
* 23601: contig of 806 bp in length
* 24407: gap of 100 bp
* 24506: contig of 823 bp in length
* 25329: gap of 100 bp
* 25430: contig of 813 bp in length
* 26243: gap of 100 bp
* 26343: contig of 812 bp in length
* 27154: gap of 100 bp
* 27255: contig of 782 bp in length
* 28037: gap of 100 bp
* 28137: contig of 801 bp in length
* 28938: gap of 100 bp
* 29038: contig of 761 bp in length
* 29799: gap of 100 bp
* 30725: contig of 827 bp in length
* 30825: gap of 100 bp

* 30826: contig of 884 bp in length
* 31709: gap of 100 bp
* 31809: contig of 812 bp in length
* 32621: gap of 100 bp
* 32721: contig of 787 bp in length
* 33508: gap of 100 bp
* 33609: contig of 792 bp in length
* 34401: gap of 100 bp
* 34501: contig of 860 bp in length
* 35361: gap of 100 bp
* 35461: contig of 824 bp in length
* 36285: gap of 100 bp
* 36385: contig of 784 bp in length
* 37169: gap of 100 bp
* 37269: contig of 799 bp in length
* 38068: gap of 100 bp
* 38167: contig of 794 bp in length
* 38961: gap of 100 bp
* 39062: contig of 815 bp in length
* 39877: gap of 100 bp
* 39977: contig of 813 bp in length
* 40790: gap of 100 bp
* 40890: contig of 852 bp in length
* 41741: gap of 100 bp
* 41842: contig of 854 bp in length
* 42695: gap of 100 bp
* 42795: contig of 870 bp in length
* 43666: gap of 100 bp
* 43766: contig of 900 bp in length
* 44666: gap of 100 bp
* 44766: contig of 896 bp in length
* 45662: gap of 100 bp
* 45762: contig of 888 bp in length
* 46499: gap of 100 bp
* 46749: contig of 872 bp in length
* 47621: gap of 100 bp
* 47721: contig of 865 bp in length
* 48587: gap of 100 bp
* 48687: contig of 852 bp in length
* 49339: gap of 100 bp
* 49639: contig of 845 bp in length
* 50484: gap of 100 bp
* 50584: contig of 855 bp in length
* 51439: gap of 100 bp
* 51539: contig of 861 bp in length
* 52400: gap of 100 bp
* 52499: contig of 889 bp in length
* 53389: gap of 100 bp
* 53489: contig of 879 bp in length
* 54368: gap of 100 bp
* 54468: contig of 877 bp in length

Query Match 73.6%; Score 36.8; DB 2; Length 118229;
Best Local Similarity 85.4%; Pred. No. 0.0031;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGCATGACTGCGCTTCCAGTCGGAACTGTGCG 50
|||||
DB 9250 TTAATTGCTGCGCTGACTGCGCTTCCAGTCGGAACTGTGCG 9297

RESULT 4

AC013368
LOCUS AC013368 211967 bp DNA linear HTG 13-JUL-2000
DEFINITION Homo sapiens clone RP11-11016, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC013368
VERSION AC013368.6 GI:9123839
KEYWORDS HTG; HTGS PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 211967)

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Birren, B., Linton, L., Nusbaum, C. and Lander, B.

Homo sapiens, Clone RP11-11016

Unpublished

2 (bases 1 to 211967)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckerly, R., Boguslavskiy, L., Boukhgalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArelano, K., Dewar, K., Domino, M., Donegan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Melidrim, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (09-NOV-1999) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 13, 2000 this sequence version replaced gi:8099785.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L3239

Center clone name: 11_O_16

* NOTE: This record contains 254 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

* 1 637: contig of 637 bp in length
* 638 737: gap of 100 bp
* 738 1380: contig of 643 bp in length
* 1381 1480: gap of 100 bp
* 1481 2102: contig of 622 bp in length
* 2103 2202: gap of 100 bp
* 2203 2824: contig of 622 bp in length
* 2825 2924: gap of 100 bp
* 2925 3591: contig of 667 bp in length
* 3592 3691: gap of 100 bp
* 3692 4364: contig of 673 bp in length
* 4365 4464: gap of 100 bp
* 4465 5118: contig of 654 bp in length
* 5119 5218: gap of 100 bp
* 5219 5879: contig of 661 bp in length
* 5880 5979: gap of 100 bp
* 5980 6655: contig of 676 bp in length
* 6656 6755: gap of 100 bp
* 6756 7382: contig of 627 bp in length
* 7383 7482: gap of 100 bp
* 7483 8135: contig of 653 bp in length
* 8136 8235: gap of 100 bp
* 8236 8900: contig of 665 bp in length
* 8901 9000: gap of 100 bp
* 9001 9652: contig of 652 bp in length
* 9653 9753: gap of 100 bp
* 9753 10378: contig of 626 bp in length
* 10379 10478: gap of 100 bp

10479 11122: contig of 644 bp in length
11123 11222: gap of 100 bp
11223 11880: contig of 658 bp in length
11881 11980: gap of 100 bp
11981 12640: contig of 660 bp in length
12641 12740: gap of 100 bp
12741 13400: contig of 660 bp in length
13401 13500: gap of 100 bp
13501 14140: contig of 640 bp in length
14141 14240: gap of 100 bp
14241 14876: contig of 636 bp in length
14877 14976: gap of 100 bp
14977 15635: contig of 659 bp in length
15636 15735: gap of 100 bp
15736 16390: contig of 655 bp in length
16391 16490: gap of 100 bp
16491 17143: contig of 653 bp in length
17144 17243: gap of 100 bp
17244 17886: contig of 643 bp in length
17887 17986: gap of 100 bp
17987 18628: contig of 642 bp in length
18629 18728: gap of 100 bp
18729 19383: contig of 655 bp in length
19384 19483: gap of 100 bp
19484 20165: contig of 682 bp in length
20166 20265: gap of 100 bp
20266 20906: contig of 641 bp in length
20907 21006: gap of 100 bp
21007 21680: contig of 674 bp in length
21681 21780: gap of 100 bp
21781 22410: contig of 630 bp in length
22411 22510: gap of 100 bp
22511 23164: contig of 654 bp in length
23165 23264: gap of 100 bp
23265 23927: contig of 663 bp in length
23928 24027: gap of 100 bp
24028 24685: contig of 658 bp in length
24686 24785: gap of 100 bp
24786 25427: contig of 642 bp in length
25428 25527: gap of 100 bp
25528 26193: contig of 666 bp in length
26194 26293: gap of 100 bp
26294 26952: contig of 659 bp in length
26953 27052: gap of 100 bp
27053 27723: contig of 671 bp in length
27724 27823: gap of 100 bp
27824 28468: contig of 645 bp in length
28469 28568: gap of 100 bp
28569 29228: contig of 660 bp in length
29229 29328: gap of 100 bp
29329 29976: contig of 648 bp in length
29977 30076: gap of 100 bp
30077 30732: contig of 656 bp in length
30733 30832: gap of 100 bp
30833 31471: contig of 639 bp in length
31472 31571: gap of 100 bp
31572 32210: contig of 639 bp in length
32211 32310: gap of 100 bp
32311 32950: contig of 640 bp in length
32951 33050: gap of 100 bp
33051 33706: contig of 656 bp in length
33707 33806: gap of 100 bp
33807 34434: contig of 628 bp in length
34435 34534: gap of 100 bp
34535 35188: contig of 654 bp in length
35189 35288: gap of 100 bp
35289 35952: contig of 664 bp in length
35953 36052: gap of 100 bp
36053 36683: contig of 631 bp in length
36684 36783: gap of 100 bp
36784 37452: contig of 669 bp in length
37453 37552: gap of 100 bp
37553 38208: contig of 656 bp in length

```

* 38209 38308: gap of 100 bp
* 38309 contig of 659 bp in length
* 38967 39067: gap of 100 bp
* 38968 39067: gap of 100 bp
* 39068 39222: contig of 655 bp in length
* 39723 39822: gap of 100 bp
* 39823 40412: contig of 590 bp in length
* 40413 40512: gap of 100 bp
* 40513 41155: contig of 643 bp in length
* 41156 41255: gap of 100 bp
* 41256 41255: contig of 670 bp in length
* 41926 42025: gap of 100 bp
* 42026 42695: contig of 670 bp in length
* 42696 42795: gap of 100 bp
* 42796 43442: contig of 647 bp in length
* 43443 43542: gap of 100 bp
* 43543 44165: contig of 623 bp in length
* 44166 44265: gap of 100 bp
* 44266 44933: contig of 668 bp in length
* 44934 45033: gap of 100 bp
* 45034 45670: contig of 637 bp in length
* 45671 45770: gap of 100 bp
* 45771 46415: contig of 645 bp in length
* 46416 46515: gap of 100 bp
* 46516 47172: contig of 657 bp in length
* 47173 47272: gap of 100 bp
* 47273 47944: contig of 672 bp in length
* 47945 48044: gap of 100 bp
* 48045 48682: contig of 638 bp in length
* 48683 48782: gap of 100 bp
* 48783 49424: contig of 642 bp in length
* 49425 49524: gap of 100 bp
* 49525 50188: contig of 664 bp in length
* 50189 50288: gap of 100 bp
* 50289 50589: contig of 670 bp in length
* 50590 51058: gap of 100 bp
* 51059 51720: contig of 662 bp in length
* 51721 51820: gap of 100 bp
* 51821 52470: contig of 650 bp in length
* 52471 52570: gap of 100 bp
* 52571 53312: contig of 742 bp in length
* 53313 53412: gap of 100 bp
* 53413 54150: contig of 738 bp in length
* 54151 54250: gap of 100 bp
* 54251 54986: contig of 736 bp in length

Query Match 73.6%; Score 36.8; DB 2; Length 211967;
Best Local Similarity 85.4%; Pred. No. 0.0032;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGCATGACTGCGCGCTTTCCAGTCGGGAACCTGTCG 50
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130366 TTAAGTCGTTGCGCTCACTGCGCGCTTTCCAGTCGGGAACCTGTCG 130413

RESULT 5
AC015852
LOCUS Homo sapiens chromosome 18 clone CTB-155C15 map 18, LOW-PASS
DEFINITION AC015852
ACCESSION AC015852
VERSION AC015852.2 GI:9144626
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 260636)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 18, clone CTB-155C15
JOURNAL Unpublished
2 (bases 1 to 260636)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhalter,B.,

```

```

Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliiev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 260636)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavkiy,L., Boukhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliiev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6446816.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L518
Center clone name: 155_C_15
-----
* NOTE: This record contains 270 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely for
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 903: contig of 903 bp in length
* 904 1003: gap of 100 bp
* 1004 1879: contig of 876 bp in length
* 1880 1979: gap of 100 bp
* 1980 2862: contig of 883 bp in length
* 2863 2962: gap of 100 bp
* 2963 3849: contig of 887 bp in length
* 3850 3949: gap of 100 bp
* 3950 4832: contig of 883 bp in length

```

* 4833 4932: gap of 100 bp
* 5810: contig of 878 bp in length
* 5811: gap of 100 bp
* 5910: gap of 100 bp
* 6899: gap of 100 bp
* 6900: gap of 100 bp
* 7777: contig of 878 bp in length
* 7778: gap of 100 bp
* 7878: gap of 100 bp
* 8750: gap of 100 bp
* 8850: gap of 100 bp
* 9727: gap of 100 bp
* 9727: gap of 100 bp
* 9727: gap of 100 bp
* 10716: gap of 100 bp
* 10815: gap of 100 bp
* 11687: contig of 872 bp in length
* 11688: gap of 100 bp
* 11788: gap of 100 bp
* 12664: contig of 877 bp in length
* 12665: gap of 100 bp
* 12665: gap of 100 bp
* 12765: contig of 880 bp in length
* 13644: contig of 880 bp in length
* 13645: gap of 100 bp
* 13745: gap of 100 bp
* 14626: contig of 882 bp in length
* 14627: gap of 100 bp
* 14727: gap of 100 bp
* 15612: contig of 886 bp in length
* 15613: gap of 100 bp
* 15713: contig of 888 bp in length
* 16600: gap of 100 bp
* 16701: gap of 100 bp
* 17572: contig of 872 bp in length
* 17573: gap of 100 bp
* 17673: gap of 100 bp
* 18554: contig of 882 bp in length
* 18555: gap of 100 bp
* 18555: contig of 879 bp in length
* 19534: gap of 100 bp
* 19534: gap of 100 bp
* 20498: contig of 865 bp in length
* 20499: gap of 100 bp
* 20599: gap of 100 bp
* 21473: contig of 875 bp in length
* 21474: gap of 100 bp
* 21573: gap of 100 bp
* 22462: contig of 889 bp in length
* 22463: gap of 100 bp
* 22562: gap of 100 bp
* 23439: contig of 877 bp in length
* 23440: gap of 100 bp
* 23539: gap of 100 bp
* 24404: contig of 865 bp in length
* 24504: gap of 100 bp
* 24505: gap of 100 bp
* 25380: contig of 876 bp in length
* 25381: gap of 100 bp
* 25480: gap of 100 bp
* 25481: contig of 892 bp in length
* 26372: gap of 100 bp
* 26373: contig of 889 bp in length
* 26473: gap of 100 bp
* 27362: gap of 100 bp
* 27461: gap of 100 bp
* 28340: contig of 879 bp in length
* 28341: gap of 100 bp
* 28440: gap of 100 bp
* 29317: contig of 877 bp in length
* 29318: gap of 100 bp
* 29417: gap of 100 bp
* 29418: contig of 886 bp in length
* 30303: contig of 886 bp in length
* 30403: gap of 100 bp
* 30404: contig of 881 bp in length
* 31285: gap of 100 bp
* 31385: contig of 871 bp in length
* 32255: gap of 100 bp
* 32256: contig of 880 bp in length
* 32356: gap of 100 bp
* 33336: gap of 100 bp
* 34222: contig of 887 bp in length
* 34322: gap of 100 bp
* 34323: contig of 875 bp in length
* 35197: gap of 100 bp
* 35198: contig of 867 bp in length
* 35298: gap of 100 bp
* 36164: gap of 100 bp
* 36165: contig of 880 bp in length
* 37144: gap of 100 bp
* 37145: gap of 100 bp
* 37245: contig of 870 bp in length
* 38115: gap of 100 bp
* 38215: contig of 890 bp in length
* 39104: gap of 100 bp
* 39204: gap of 100 bp
* 40089: contig of 885 bp in length
* 40189: gap of 100 bp

* 40190 41061: contig of 872 bp in length
* 41062 41161: gap of 100 bp
* 41162 42037: contig of 876 bp in length
* 42038 42137: gap of 100 bp
* 42138 43012: contig of 875 bp in length
* 43013 43112: gap of 100 bp
* 43113 43980: contig of 868 bp in length
* 43981 44080: gap of 100 bp
* 44081 44959: contig of 879 bp in length
* 44960 45059: gap of 100 bp
* 45060 45943: contig of 884 bp in length
* 45944 46043: gap of 100 bp
* 46044 46917: contig of 874 bp in length
* 46918 47017: gap of 100 bp
* 47018 47894: contig of 877 bp in length
* 47895 48870: contig of 876 bp in length
* 48871 48970: gap of 100 bp
* 48971 49844: contig of 874 bp in length
* 49845 49944: gap of 100 bp
* 49945 50826: contig of 882 bp in length
* 50827 50926: gap of 100 bp
* 50927 51806: contig of 880 bp in length
* 51807 51906: gap of 100 bp
* 51907 52777: contig of 871 bp in length
* 52778 52877: gap of 100 bp
* 52878 53760: contig of 883 bp in length
* 53761 53860: gap of 100 bp
* 53861 54727: contig of 867 bp in length
* 54728 54827: gap of 100 bp
* 54828 55699: contig of 872 bp in length
* 55700 55799: gap of 100 bp
* 55800 56689: contig of 890 bp in length
* 56690 56789: gap of 100 bp
* 56790 57663: contig of 874 bp in length
* 57664 57763: gap of 100 bp
* 57764 58633: contig of 870 bp in length

Query Match 73.6%; Score 36.8; DB 2; Length 260636;
Best Local Similarity 85.4%; Pred. No. 0.0032;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGCGATGACTGCCGCTTTCCAGTCGGAAACCTGCG 50
|||||
Db 128125 TTAATTGTGTGCGCTCACTGCCGCTTTCCAGTCGGAAACCTGCG 128172

RESULT 6
AC010800

LOCUS Homo sapiens clone 1_E_17, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC010800 73282 bp DNA linear HTG 23-SEP-1999

ACCESSION AC010800

VERSION AC010800.1 GI:5919335

KEYWORDS HTG; HTGS. PHASE0.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 73282)

AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE Homo sapiens, clone 1_E_17

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 73282)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Balwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Collings,S., Collins,A., Collymore,A.,
Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donnell,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tsefaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

TITLE
JOURNAL

Direct Submission
Submitted (23-SEP-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker: Smit, A.F.A. &
Green, P. (1996-1997)

COMMENT

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.

* NOTE: This record contains 86 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

* 1 859: contig of 859 bp in length

* gap of unknown length

* 860 1695: contig of 836 bp in length

* gap of unknown length

* 1696 2546: contig of 851 bp in length

* gap of unknown length

* 2547 3398: contig of 852 bp in length

* gap of unknown length

* 3399 4263: contig of 865 bp in length

* gap of unknown length

* 4264 5163: contig of 900 bp in length

* gap of unknown length

* 5164 6030: contig of 867 bp in length

* gap of unknown length

* 6031 6901: contig of 871 bp in length

* gap of unknown length

* 6902 7717: contig of 816 bp in length

* gap of unknown length

* 7718 8565: contig of 848 bp in length

* gap of unknown length

* 8566 9412: contig of 847 bp in length

* gap of unknown length

* 9413 10266: contig of 854 bp in length

* gap of unknown length

* 10267 11243: contig of 977 bp in length

* gap of unknown length

* 11244 12098: contig of 855 bp in length

* gap of unknown length

* 12099 12940: contig of 842 bp in length

* gap of unknown length

* 12941 13799: contig of 859 bp in length

* gap of unknown length

* 13800 14670: contig of 871 bp in length

* gap of unknown length

* 14671 15489: contig of 819 bp in length

* gap of unknown length

* 15490 16379: contig of 890 bp in length

* gap of unknown length

* 16380 17216: contig of 837 bp in length

* gap of unknown length

* 17217 18204: contig of 988 bp in length

* gap of unknown length

* 18205 19021: contig of 817 bp in length

* gap of unknown length

* 19022 19877: contig of 856 bp in length

* gap of unknown length

* 19878 20718: contig of 841 bp in length

* gap of unknown length

* 20719 21555: contig of 837 bp in length

* gap of unknown length

* 21556 22372: contig of 817 bp in length

* gap of unknown length

* 22373 23227: contig of 855 bp in length

* gap of unknown length

* 23228 24050: contig of 823 bp in length

* gap of unknown length

* 24051 24905: contig of 855 bp in length

* gap of unknown length

* 24906 25754: contig of 849 bp in length

* gap of unknown length

* 25755 26583: contig of 829 bp in length

* gap of unknown length

* 26584 27409: contig of 826 bp in length

* gap of unknown length

* 27410 28215: contig of 806 bp in length

* gap of unknown length

* 28216 29040: contig of 825 bp in length

* gap of unknown length

* 29041 29893: contig of 853 bp in length

* gap of unknown length

* 29894 30736: contig of 843 bp in length

* gap of unknown length

* 30737 31575: contig of 839 bp in length

* gap of unknown length

* 31576 32422: contig of 847 bp in length

* gap of unknown length

* 32423 33293: contig of 871 bp in length

* gap of unknown length

* 33294 34134: contig of 841 bp in length

* gap of unknown length

* 34135 34971: contig of 837 bp in length

* gap of unknown length

* 34972 35789: contig of 818 bp in length

* gap of unknown length

* 35790 36632: contig of 843 bp in length

* gap of unknown length

* 36633 37478: contig of 846 bp in length

* gap of unknown length

* 37479 38321: contig of 843 bp in length

* gap of unknown length

* 38322 39163: contig of 842 bp in length

* gap of unknown length

* 39164 40013: contig of 850 bp in length

* gap of unknown length

* 40014 40858: contig of 845 bp in length

* gap of unknown length

* 40859 41719: contig of 861 bp in length

* gap of unknown length

* 41720 42581: contig of 862 bp in length

* gap of unknown length

* 42582 43449: contig of 868 bp in length

* gap of unknown length

* 43450 44290: contig of 841 bp in length

* gap of unknown length

* 44291 45216: contig of 926 bp in length

* gap of unknown length

* 45217 46065: contig of 849 bp in length

* gap of unknown length

* 46066 46890: contig of 825 bp in length

* gap of unknown length

* 46891 47747: contig of 857 bp in length

* gap of unknown length

* 47748 48582: contig of 835 bp in length

* gap of unknown length

* 48583 49444: contig of 862 bp in length

* gap of unknown length

* 49445 50302: contig of 858 bp in length

* gap of unknown length

* 50303 51154: contig of 852 bp in length

* gap of unknown length

* 51155 52011: contig of 857 bp in length

* gap of unknown length

* 52012 52864: contig of 853 bp in length

* gap of unknown length

* 52865 53680: contig of 816 bp in length

```

* 53681 gap of unknown length
* 54500: contig of 820 bp in length
* 54501 gap of unknown length
* 55357: contig of 857 bp in length
* 55358 gap of unknown length
* 56213: contig of 856 bp in length
* 57033: gap of unknown length
* 57034: contig of 820 bp in length
* 58052: gap of unknown length
* 58053: contig of 1019 bp in length
* 58054: gap of unknown length
* 58884: contig of 832 bp in length
* 58885 gap of unknown length
* 59738: contig of 854 bp in length
* 59739 gap of unknown length
* 60599: contig of 861 bp in length
* 60600 gap of unknown length
* 61442: contig of 843 bp in length
* 61443 gap of unknown length
* 62284: contig of 842 bp in length
* 62285 gap of unknown length
* 63113: contig of 829 bp in length
* 63114 gap of unknown length
* 63949: contig of 836 bp in length
* 63950 gap of unknown length
* 64788: contig of 839 bp in length
* 64789 gap of unknown length
* 65709: contig of 921 bp in length
* 65710 gap of unknown length
* 6552: contig of 843 bp in length

Query Match 72.4%; Score 36.2; DB 2; Length 73282;
Best Local Similarity 82.0%; Pred. No. 0.0053;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ACTTACCTTATAGCATGCTCCGCTTTCAGTCGGAAACCTGTCG 50
DB 59421 ANTAAATGCTTGTGCTACTGCTCCGCTTTCAGTCGGAAACCTGTCG 59470

RESULT 7
AC073998
LOCUS AC073998 217141 bp DNA linear HTG 30-SBP-2000
DEFINITION Homo sapiens clone RP11-337M23, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC073998
VERSION AC073998.4 GI:10440718
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (base 1 to 217141)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-337M23
Unpublished
2 (bases 1 to 217141)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Bogulavskiy,L., Boukhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choelel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Menes,L., Mihova,T., Miranda,C., Milenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,

```

Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (09-JUL-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Sep 30, 2000 this sequence version replaced gi:9838018.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L7934

Center clone name: 337_M_23

* NOTE: This record contains 270 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

```

* 1 685: contig of 685 bp in length
* 686 785: gap of 100 bp
* 786 1466: contig of 681 bp in length
* 1467 1566: gap of 100 bp
* 1567 2244: contig of 678 bp in length
* 2245 2344: gap of 100 bp
* 2345 3018: contig of 674 bp in length
* 3019 3118: gap of 100 bp
* 3119 3779: contig of 661 bp in length
* 3780 3879: gap of 100 bp
* 3880 4565: contig of 686 bp in length
* 4566 4665: gap of 100 bp
* 4666 5345: contig of 680 bp in length
* 5346 5445: gap of 100 bp
* 5446 6122: contig of 677 bp in length
* 6123 6222: gap of 100 bp
* 6223 6891: contig of 669 bp in length
* 6892 6992: contig of 100 bp
* 6993 7674: contig of 683 bp in length
* 7675 7774: gap of 100 bp
* 7775 8458: contig of 684 bp in length
* 8459 8558: gap of 100 bp
* 8559 9249: contig of 691 bp in length
* 9250 9349: gap of 100 bp
* 9350 10031: contig of 682 bp in length
* 10032 10131: gap of 100 bp
* 10132 10831: contig of 700 bp in length
* 10832 10931: gap of 100 bp
* 10932 11617: contig of 686 bp in length
* 11618 11717: gap of 100 bp
* 11718 12390: contig of 673 bp in length
* 12391 12490: gap of 100 bp
* 12491 13169: contig of 679 bp in length
* 13170 13269: gap of 100 bp
* 13270 13947: contig of 678 bp in length
* 13948 14047: gap of 100 bp
* 14048 14714: contig of 667 bp in length
* 14715 14814: gap of 100 bp
* 14815 15505: contig of 691 bp in length
* 15506 15605: gap of 100 bp
* 15606 16284: contig of 679 bp in length
* 16285 16384: gap of 100 bp

```

* 16385 17071: contig of 687 bp in length
* 17072 17171: gap of 100 bp
* 17172 17837: contig of 666 bp in length
* 17838 17937: gap of 100 bp
* 17938 18614: contig of 677 bp in length
* 18615 18714: gap of 100 bp
* 18715 19396: contig of 682 bp in length
* 19397 19496: gap of 100 bp
* 19497 20186: contig of 690 bp in length
* 20187 20286: gap of 100 bp
* 20287 20970: contig of 684 bp in length
* 20971 21070: gap of 100 bp
* 21071 21757: contig of 687 bp in length
* 21758 21857: gap of 100 bp
* 21858 22550: contig of 693 bp in length
* 22551 22650: gap of 100 bp
* 22651 23332: contig of 682 bp in length
* 23333 23432: gap of 100 bp
* 23433 24115: contig of 683 bp in length
* 24116 24215: gap of 100 bp
* 24216 24899: contig of 684 bp in length
* 24900 24999: gap of 100 bp
* 25000 25682: contig of 683 bp in length
* 25683 25782: gap of 100 bp
* 25783 26474: contig of 692 bp in length
* 26475 26574: gap of 100 bp
* 26575 27258: contig of 684 bp in length
* 27259 27358: gap of 100 bp
* 27359 28040: contig of 682 bp in length
* 28041 28140: gap of 100 bp
* 28141 28921: contig of 681 bp in length
* 28922 28921: gap of 100 bp
* 28922 29601: contig of 680 bp in length
* 29602 29701: gap of 100 bp
* 29702 30385: contig of 684 bp in length
* 30386 30485: gap of 100 bp
* 30486 31173: contig of 688 bp in length
* 31174 31273: gap of 100 bp
* 31274 31959: contig of 686 bp in length
* 31960 32059: gap of 100 bp
* 32060 32742: contig of 683 bp in length
* 32743 32842: gap of 100 bp
* 32843 33528: contig of 686 bp in length
* 33529 33629: gap of 100 bp
* 33629 34221: contig of 693 bp in length
* 34222 34421: gap of 100 bp
* 34422 35119: contig of 698 bp in length
* 35120 35219: gap of 100 bp
* 35220 35998: contig of 679 bp in length
* 35999 35998: gap of 100 bp
* 35999 36682: contig of 684 bp in length
* 36683 36782: gap of 100 bp
* 36783 37470: contig of 688 bp in length
* 37471 37570: gap of 100 bp
* 37571 38240: contig of 670 bp in length
* 38241 38340: gap of 100 bp
* 38341 39025: contig of 685 bp in length
* 39026 39125: gap of 100 bp
* 39126 39824: contig of 699 bp in length
* 39825 39924: gap of 100 bp
* 39925 40610: contig of 686 bp in length
* 40611 40710: gap of 100 bp
* 40711 41397: contig of 687 bp in length
* 41398 41497: gap of 100 bp
* 41498 42184: contig of 687 bp in length
* 42185 42284: gap of 100 bp
* 42285 42968: contig of 684 bp in length
* 42969 43068: gap of 100 bp
* 43069 43759: contig of 691 bp in length
* 43760 43859: gap of 100 bp
* 43860 44530: contig of 671 bp in length
* 44531 44630: gap of 100 bp
* 44631 45318: contig of 688 bp in length

* 45319 45418: gap of 100 bp
* 45419 46098: contig of 680 bp in length
* 46099 46198: gap of 100 bp
* 46199 46860: contig of 662 bp in length
* 46861 46960: gap of 100 bp
* 46961 47650: contig of 690 bp in length
* 47651 47750: gap of 100 bp
* 47751 48425: contig of 675 bp in length
* 48426 48525: gap of 100 bp
* 48526 49210: contig of 685 bp in length
* 49211 49310: gap of 100 bp
* 49311 49997: contig of 687 bp in length
* 49998 50097: gap of 100 bp
* 50098 50790: contig of 693 bp in length
* 50791 50890: gap of 100 bp
* 50891 51580: contig of 690 bp in length
* 51581 51680: gap of 100 bp
* 51681 52361: contig of 681 bp in length
* 52362 52461: gap of 100 bp
* 52462 53117: contig of 656 bp in length
* 53118 53217: gap of 100 bp
* 53218 53903: contig of 686 bp in length
* 53904 54003: gap of 100 bp

Query Match 72.4%; Score 36.2; DB 2; Length 217141;

Best Local Similarity 83.7%; Pred. No. 0.0054; Mismatches 41; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 CTTAAGCTTATAGCGATGATGCCCGCTTTCACGTCGGGAAACCTGCG 50
DB 162194 CTTAATTCGTTGGCTCACTGCCCGCTTTCACGTCGGGAAACCTGCG 162242

RESULT 8

G39061

LOCUS

DEFINITION Z11905 Zebrafish AB Danio rerio STS genomic, sequence tagged site.

ACCESSION G39061

VERSION G39061.1 GI:3358270

KEYWORDS STS.

SOURCE Danio rerio (zebrafish)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 604)

Shimoda, N., Knapik, E.W., Ziniti, J., Sim, C., Yamada, E., Kaplan, S.,

Jackson, D., de Sauvage, F., Jacob, H. and Fishman, M.C.

Zebrafish genetic map with 2000 microsatellite markers

Genomics 58 (3), 219-232 (1999)

99303552

STS size: 112

PCR Profile:

Presoak: 94 degrees C for 5.0 minutes

Denaturation: 94 degrees C for 1.0 minute

Annealing: 58 degrees C for 1.0 minute

Polymerization: 72 degrees C for 1.5 minute

PCR Cycles: 27

Thermal Cycler: MJ Research PTC-100

Protocol:

Template: 10 ng

Primer: each 375 nM

Contact: Mark C. Fishman

Cardiovascular Research Center

Massachusetts General Hospital

Mail code 1494100A, 149 13th Street, Charlestown, MA 02129, USA

Fax: 617265806

Email: fishman@gh.cvr.harvard.edu

http://zebrafish.mgh.harvard.edu

Primer A: CACGAGCTTCACTGACGTA

Primer B: ATACACACCCAGCCGACAT

Primer B: TCTCTCCCTGGACATCATC
STS size: 142
PCR Profile:

Presoak: 94 degrees C for 5.0 minutes
Denaturation: 94 degrees C for 1.0 minute
Annealing: 58 degrees C for 1.0 minute
Polymerization: 72 degrees C for 1.5 minute
PCR Cycles: 27
Thermal Cycler: MJ Research PTC-100

Protocol:

Template: 10 ng
Primer: each 375 nM
dNTPs: each 200 uM
Tag Polymerase: 0.034 units/ul
Total Vol: 10 ul

Buffer: MgCL2: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3

Primers are available from Research Genetics Inc.
(http://www.resgen.com phone: 800-533-4363).

FEATURES

Location/Qualifiers
1..826
/organism="Danio rerio"
/mol_type="genomic DNA"
/strain="AB"
/db_xref="taxon:7955"
/map="LG 16"
/sex="F"
/clone_lib="Zebrafish AB"
/dev_stage="Adult"
/lab_host="DH5alphaP' IQ"
/note="Vector: m13mp19 with added BstXI site; V-type: Phage; Genomic DNA from a single adult Zebrafish of AB strain was digested with AluI, Cac8I, HaeIII, NlaVI, or RsaI. Fragments in the range of 250-500 bp were gel purified and a BstXI linker was added. The fragments were cloned into a modified M13mp19 vector and transformed into E. Coli DH5alpha. Microsatellite sequences were screened with labeled d(CA)15 and d(GT)15 oligonucleotide probes."
65..206
65..84
complement(187..206)

STS
primer_bind
primer_bind
ORIGIN

Query Match 71.6%; Score 35.8; DB 11; Length 826;
Best Local Similarity 83.3%; Pred. No. 0.0068;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGCGATGCTGCCCGCTTCAGTCGGGAACCTGTGCG 50
|||||
Db 516 TTAANTGCGTTCGCTCACTGCCGCTTCAGTCGGGAACCTGTGCG 563
|||||

RESULT 11

G40162 979 bp DNA linear STS 01-FEB-2001
LOCUS
DEFINITION
220177 Zebrafish AB Danio rerio STS genomic, sequence tagged site.

ACCESSION G40162

VERSION G40162.1 GI:3359371

KEYWORDS STS.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 979)

REFERENCE

AUTHORS Shimoda,N., Knapik,E.W., Ziniti,J., Sim,C., Yamada,E., Kaplan,S.,
Jackson,D., de Sauvage,F., Jacob,H. and Fishman,M.C.

TITLE JOURNAL MEDLINE PUBMED COMMENT

Zebrafish genetic map with 2000 microsatellite markers
Genomics 58 (3), 219-232 (1999)
99303552
10373319

Contact: Mark C. Fishman
Cardiovascular Research Center
Massachusetts General Hospital
Mail code 1494100A, 149 13th Street, Charlestown, MA 02129, USA
Fax: 6177265806
Email: fishman@gh.cvr.harvard.edu
http://zebrafish.mgh.harvard.edu
Primer A: TCAATCTGCAAACTCCGCA
Primer B: CGCTTATAGGCTGCAGAG
STS size: 244
PCR Profile:

Presoak: 94 degrees C for 5.0 minutes
Denaturation: 94 degrees C for 1.0 minute
Annealing: 58 degrees C for 1.0 minute
Polymerization: 72 degrees C for 1.5 minute
PCR Cycles: 27
Thermal Cycler: MJ Research PTC-100

Protocol:

Template: 10 ng
Primer: each 375 nM
dNTPs: each 200 uM
Tag Polymerase: 0.034 units/ul
Total Vol: 10 ul

Buffer:

MgCL2: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3

Primers are available from Research Genetics Inc.
(http://www.resgen.com phone: 800-533-4363).

FEATURES

Location/Qualifiers
1..979
/organism="Danio rerio"
/mol_type="genomic DNA"
/strain="AB"
/db_xref="taxon:7955"
/map="LG 16"
/sex="F"
/clone_lib="Zebrafish AB"
/dev_stage="Adult"
/lab_host="DH5alphaP' IQ"
/note="Vector: m13mp19 with added BstXI site; V-type: Phage; Genomic DNA from a single adult Zebrafish of AB strain was digested with AluI, Cac8I, HaeIII, NlaVI, or RsaI. Fragments in the range of 250-500 bp were gel purified and a BstXI linker was added. The fragments were cloned into a modified M13mp19 vector and transformed into E. Coli DH5alpha. Microsatellite sequences were screened with labeled d(CA)15 and d(GT)15 oligonucleotide probes."
120..363
120..139
complement(344..363)

STS
primer_bind
primer_bind
ORIGIN

Query Match 71.6%; Score 35.8; DB 11; Length 979;
Best Local Similarity 83.3%; Pred. No. 0.0068;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGCGATGCTGCCCGCTTCAGTCGGGAACCTGTGCG 50
|||||
Db 726 TTAANTGCGTTCGCTCACTGCCGCTTCAGTCGGGAACCTGTGCG 773
|||||

RESULT 12
AC023542

```

LOCUS      AC023542      76295 bp      DNA      linear      HTG 13-JUL-2000
DEFINITION Homo sapiens chromosome 15 clone RP11-165P21 map 15, LOW-PASS
SEQUENCE SAMPLING.
ACCESSION  AC023542
VERSION    AC023542.2 GI:9156023
KEYWORDS   HTG; HTGS PHASE0
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 76295)
AUTHORS   Birren B., Linton L., Nusbaum C. and Lander E.
TITLE     Homo sapiens chromosome 15, clone RP11-165P21
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 76295)
AUTHORS   Birren B., Linton L., Nusbaum C., Lander E., Abraham H., Allen N.,
            Anderson S., Baldwin J., Barna N., Beda P., Boguslavsky L.,
            Boukhalter B., Brown A., Burkett G., Campopiano A., Castle A.,
            Choepel Y., Collangelo M., Collins S., Collamore A., Cooke P.,
            DeArelano K., Dewar K., Dodge S., Domino M., Doyle M.,
            Feneator J., Ferreira P., FitzHugh W., Forrest C., Gage D.,
            Galagan J., Gardyna S., Ginde S., Goyette M., Graham L.,
            Grand-Pierre N., Grant G., Hagos B., Heaford A., Horton L.,
            Howland J.C., Iliev I., Johnson R., Jones C., Kann L., Karatas A.,
            Klein J., Landers T., Largocque K., Lehoczy J., Levine R.,
            Liu C., Liu G., Locke K., Macdonald P., Marquis N., McCarthy M.,
            McEwan P., McGurk A., McKernan K., McPheeters R., Meldrim J.,
            Meneus L., Mihova T., Miranda C., Mlenga V., Morrow J., Naylor J.,
            Norman C.H., O'Connor T., O'Donnell P., O'Neil D., Olivari T.M.,
            Peterson K., Pierre N., Pisani C., Pollara V., Raymond C.,
            Riley R., Rogov P., Rothman D., Roy A., Santos R., Schauer S.,
            Severy P., Spencer B., Stange-Thomann N., Stojanovic N.,
            Subramanian A., Talamas J., Tesfaye S., Theodore J., Tirrell A.,
            Travers M., Trigilio J., Vassiliev H., Viel R., Vo A., Wilson B.,
            Wu X., Wyman D., Ye W.J., Young G., Zainoun J., Zimmer A. and
            Zody M.
DIRECT SUBMISSION
Submitted (15-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6978228.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L6028
Center clone name: 165_P_21
-----
* NOTE: This record contains 86 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1      788: contig of 788 bp in length
*      889: gap of 100 bp
*      889: contig of 783 bp in length
*      1672: gap of 100 bp
*      1771: gap of 100 bp
*      1772: contig of 792 bp in length
*      2563: contig of 792 bp in length
*      2564: gap of 100 bp
*      2664: contig of 791 bp in length
*      3454: contig of 791 bp in length
*      3455: gap of 100 bp
*      4331: contig of 777 bp in length
*      4332: gap of 100 bp
*
* 4432: contig of 827 bp in length
* 5258: gap of 100 bp
* 5358: contig of 796 bp in length
* 6154: gap of 100 bp
* 6254: gap of 100 bp
* 6255: contig of 788 bp in length
* 7043: gap of 100 bp
* 7142: contig of 777 bp in length
* 7919: gap of 100 bp
* 8019: gap of 100 bp
* 8788: contig of 769 bp in length
* 8888: gap of 100 bp
* 8889: contig of 779 bp in length
* 9667: gap of 100 bp
* 9668: contig of 799 bp in length
* 10567: gap of 100 bp
* 10567: contig of 778 bp in length
* 11444: gap of 100 bp
* 11445: contig of 777 bp in length
* 12321: gap of 100 bp
* 12322: contig of 779 bp in length
* 12421: gap of 100 bp
* 13200: contig of 779 bp in length
* 13201: gap of 100 bp
* 14086: contig of 786 bp in length
* 14087: gap of 100 bp
* 14186: contig of 798 bp in length
* 14984: gap of 100 bp
* 14985: contig of 775 bp in length
* 15085: gap of 100 bp
* 15846: gap of 100 bp
* 15847: contig of 780 bp in length
* 15947: gap of 100 bp
* 16727: contig of 796 bp in length
* 16727: gap of 100 bp
* 16827: contig of 760 bp in length
* 17587: gap of 100 bp
* 17587: contig of 775 bp in length
* 17687: gap of 100 bp
* 18462: contig of 779 bp in length
* 18562: gap of 100 bp
* 19341: contig of 796 bp in length
* 19440: gap of 100 bp
* 20336: contig of 796 bp in length
* 20337: gap of 100 bp
* 20337: contig of 783 bp in length
* 21119: gap of 100 bp
* 21219: contig of 796 bp in length
* 22015: gap of 100 bp
* 22115: gap of 100 bp
* 22919: contig of 804 bp in length
* 23019: gap of 100 bp
* 23816: contig of 797 bp in length
* 23916: gap of 100 bp
* 24788: contig of 772 bp in length
* 24788: gap of 100 bp
* 25596: contig of 808 bp in length
* 25596: gap of 100 bp
* 26481: contig of 785 bp in length
* 26581: gap of 100 bp
* 27368: contig of 787 bp in length
* 27468: gap of 100 bp
* 28245: contig of 777 bp in length
* 28345: gap of 100 bp
* 29118: contig of 773 bp in length
* 29218: gap of 100 bp
* 30006: contig of 788 bp in length
* 30106: gap of 100 bp
* 30897: contig of 791 bp in length
* 30997: gap of 100 bp
* 31778: contig of 781 bp in length
* 31878: gap of 100 bp
* 32652: contig of 774 bp in length
* 32752: gap of 100 bp
* 33532: contig of 780 bp in length
* 33632: gap of 100 bp
* 34408: contig of 776 bp in length
* 34508: gap of 100 bp
* 35276: contig of 768 bp in length
* 35376: gap of 100 bp
* 36155: contig of 779 bp in length
* 36255: gap of 100 bp
* 37048: contig of 793 bp in length

```


Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center Project name: L5114
 Center Clone name: 589_i_12

* NOTE: This record contains 88 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 724: contig of 724 bp in length
 * 725 824: gap of 100 bp
 * 825 1573: contig of 749 bp in length
 * 1574 1673: gap of 100 bp
 * 1674 2455: contig of 782 bp in length
 * 2456 2555: gap of 100 bp
 * 2556 3328: contig of 773 bp in length
 * 3329 3428: gap of 100 bp
 * 3429 4179: contig of 751 bp in length
 * 4180 4279: gap of 100 bp
 * 4280 5022: contig of 743 bp in length
 * 5023 5122: gap of 100 bp
 * 5123 5855: contig of 743 bp in length
 * 5856 5965: gap of 100 bp
 * 5966 6703: contig of 738 bp in length
 * 6704 6803: gap of 100 bp
 * 6804 7554: contig of 751 bp in length
 * 7555 7654: gap of 100 bp
 * 7655 8408: contig of 754 bp in length
 * 8409 8508: gap of 100 bp
 * 8509 9254: contig of 746 bp in length
 * 9255 9354: gap of 100 bp
 * 9355 10118: contig of 764 bp in length
 * 10119 10218: gap of 100 bp
 * 10219 10978: contig of 760 bp in length
 * 10979 11078: gap of 100 bp
 * 11079 11850: contig of 772 bp in length
 * 11851 11950: gap of 100 bp
 * 11951 12712: contig of 762 bp in length
 * 12713 12812: gap of 100 bp
 * 12813 13553: contig of 741 bp in length
 * 13554 13653: gap of 100 bp
 * 13654 14392: contig of 739 bp in length
 * 14393 14492: gap of 100 bp
 * 14493 15244: contig of 752 bp in length
 * 15245 15344: gap of 100 bp
 * 15345 16095: contig of 751 bp in length
 * 16096 16195: gap of 100 bp
 * 16196 16927: contig of 732 bp in length
 * 16928 17027: gap of 100 bp
 * 17028 17766: contig of 739 bp in length
 * 17767 17866: gap of 100 bp
 * 17867 18615: contig of 749 bp in length
 * 18616 18715: gap of 100 bp
 * 18716 19489: contig of 774 bp in length
 * 19490 19589: gap of 100 bp
 * 19590 20344: contig of 755 bp in length
 * 20345 20444: gap of 100 bp
 * 20445 21205: contig of 761 bp in length
 * 21206 21305: gap of 100 bp
 * 21306 22089: contig of 784 bp in length
 * 22090 22189: gap of 100 bp
 * 22190 22972: contig of 783 bp in length
 * 22973 23072: gap of 100 bp
 * 23073 23826: contig of 754 bp in length
 * 23827 23926: gap of 100 bp
 * 23927 24674: contig of 748 bp in length

* 24675 24774: gap of 100 bp
 * 24775 25507: contig of 733 bp in length
 * 25508 25607: gap of 100 bp
 * 25608 26347: contig of 740 bp in length
 * 26348 26447: gap of 100 bp
 * 26448 27202: contig of 755 bp in length
 * 27203 27302: gap of 100 bp
 * 27303 28052: contig of 750 bp in length
 * 28053 28152: gap of 100 bp
 * 28153 28909: contig of 757 bp in length
 * 28910 29009: gap of 100 bp
 * 29010 29741: contig of 732 bp in length
 * 29742 29841: gap of 100 bp
 * 29842 30624: contig of 783 bp in length
 * 30625 30724: gap of 100 bp
 * 30725 31497: contig of 773 bp in length
 * 31498 31597: gap of 100 bp
 * 31598 32373: contig of 776 bp in length
 * 32374 32473: gap of 100 bp
 * 32474 33222: contig of 749 bp in length
 * 33223 33322: gap of 100 bp
 * 33323 34072: contig of 750 bp in length
 * 34073 34172: gap of 100 bp
 * 34173 34932: contig of 760 bp in length
 * 34933 35032: gap of 100 bp
 * 35033 35791: contig of 759 bp in length
 * 35792 35891: gap of 100 bp
 * 35892 36636: contig of 745 bp in length
 * 36637 36736: gap of 100 bp
 * 36737 37468: contig of 732 bp in length
 * 37469 37568: gap of 100 bp
 * 37569 38328: contig of 760 bp in length
 * 38329 38428: gap of 100 bp
 * 38429 39191: contig of 763 bp in length
 * 39192 39291: gap of 100 bp
 * 39292 40036: contig of 745 bp in length
 * 40037 40136: gap of 100 bp
 * 40137 40895: contig of 759 bp in length
 * 40896 40995: gap of 100 bp
 * 40996 41729: contig of 734 bp in length
 * 41730 41829: gap of 100 bp
 * 41830 42577: contig of 748 bp in length
 * 42578 42677: gap of 100 bp
 * 42678 43414: contig of 737 bp in length
 * 43415 43514: gap of 100 bp
 * 43515 44267: contig of 753 bp in length
 * 44268 44367: gap of 100 bp
 * 44368 45123: contig of 756 bp in length
 * 45124 45223: gap of 100 bp
 * 45224 45973: contig of 750 bp in length
 * 45974 46073: gap of 100 bp
 * 46074 46816: contig of 743 bp in length
 * 46817 46916: gap of 100 bp
 * 46917 47681: contig of 765 bp in length
 * 47682 47781: gap of 100 bp
 * 47782 48544: contig of 763 bp in length
 * 48545 48644: gap of 100 bp
 * 48645 49410: contig of 766 bp in length
 * 49411 49510: gap of 100 bp
 * 49511 50247: contig of 737 bp in length
 * 50248 50347: gap of 100 bp
 * 50348 51067: contig of 720 bp in length
 * 51068 51167: gap of 100 bp
 * 51168 51915: contig of 748 bp in length
 * 51916 52015: gap of 100 bp
 * 52016 52745: contig of 730 bp in length
 * 52746 52845: gap of 100 bp
 * 52846 53584: contig of 739 bp in length
 * 53585 53684: gap of 100 bp
 * 53685 54427: contig of 743 bp in length
 * 54428 54527: gap of 100 bp
 * 54528 55292: contig of 765 bp in length
 * 55293 55392: gap of 100 bp

```

* 55393 56164: contig of 772 bp in length
* 56165 56264: gap of 100 bp
* 56265 56985: contig of 721 bp in length
* 56986 57086: gap of 100 bp
* 57086 57829: contig of 744 bp in length
* 57830 57929: gap of 100 bp
* 57930 58677: contig of 748 bp in length

Query Match
Best Local Similarity 71.2%; Score 35.6; DB 2; Length 75002;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ACTTAAGCTTATAGCGATGACGCGGCTTCCAGTCGGGAACCTGCG 50
    |||||
Db 25456 AATTAAATGCGTGGCTCAGTCGCGGCTTCCAGTCGGGAACCTGCG 25505
    |||||

RESULT 15
AC016798
LOCUS
DEFINITION
AC016798
AC016798.2 GI:9119829
VERSION
HTG; HTGS PHASE0.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 95127)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Bouckghalter,B.,
Brown,A., Castle,A., Collangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meidrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (07-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6532117.
All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2499
Center clone name: 1_J_15
-----
* NOTE: This record contains 97 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will

```

```

* be preserved.
* 1 903: contig of 903 bp in length
* 904 1003: gap of 100 bp
* 1004 1852: contig of 849 bp in length
* 1853 2797: contig of 100 bp
* 2798 3777: contig of 880 bp in length
* 3778 4737: contig of 860 bp in length
* 4738 5670: contig of 833 bp in length
* 5671 6639: contig of 859 bp in length
* 6640 7619: contig of 880 bp in length
* 7620 8600: contig of 881 bp in length
* 8601 9570: contig of 870 bp in length
* 9571 10521: contig of 851 bp in length
* 10522 11452: contig of 831 bp in length
* 11453 12420: contig of 868 bp in length
* 12421 13520: contig of 870 bp in length
* 13521 13490: gap of 100 bp
* 13491 14361: contig of 871 bp in length
* 14362 15330: contig of 869 bp in length
* 15331 16430: gap of 100 bp
* 16431 17312: contig of 882 bp in length
* 17313 17412: gap of 100 bp
* 17413 18298: contig of 886 bp in length
* 18299 19243: contig of 845 bp in length
* 19244 20190: contig of 847 bp in length
* 20191 21162: contig of 872 bp in length
* 21163 21262: gap of 100 bp
* 21263 22130: contig of 868 bp in length
* 22131 23094: contig of 864 bp in length
* 23095 23194: gap of 100 bp
* 23195 24072: contig of 878 bp in length
* 24073 25016: contig of 844 bp in length
* 25017 25993: contig of 877 bp in length
* 25994 26980: contig of 887 bp in length
* 26981 27080: gap of 100 bp
* 27081 27903: contig of 829 bp in length
* 27904 28886: contig of 877 bp in length
* 28887 29834: contig of 848 bp in length
* 29835 30798: contig of 864 bp in length
* 30799 31768: contig of 870 bp in length
* 31769 32765: contig of 897 bp in length
* 32766 33666: contig of 801 bp in length
* 33667 34665: contig of 899 bp in length
* 34666 34765: gap of 100 bp

```

* 34766 35590: contig of 825 bp in length
* 35591 35690: gap of 100 bp
* 35691 36528: contig of 838 bp in length
* 36529 36628: gap of 100 bp
* 36629 37489: contig of 861 bp in length
* 37490 37589: gap of 100 bp
* 37590 38458: contig of 879 bp in length
* 38459 38568: gap of 100 bp
* 38569 39546: contig of 878 bp in length
* 39547 39546: gap of 100 bp
* 40373 40472: contig of 826 bp in length
* 40473 41361: contig of 889 bp in length
* 41362 41461: gap of 100 bp
* 41462 42333: contig of 872 bp in length
* 42334 42433: gap of 100 bp
* 42434 43313: contig of 880 bp in length
* 43314 43413: gap of 100 bp
* 43414 44242: contig of 829 bp in length
* 44243 44342: gap of 100 bp
* 44343 45211: contig of 869 bp in length
* 45212 45311: gap of 100 bp
* 45312 46183: contig of 872 bp in length
* 46184 46283: gap of 100 bp
* 46284 47132: contig of 849 bp in length
* 47133 47232: gap of 100 bp
* 47233 48077: contig of 845 bp in length
* 48078 48177: gap of 100 bp
* 48178 49045: contig of 868 bp in length
* 49046 49145: gap of 100 bp
* 49146 50014: contig of 869 bp in length
* 50015 50114: gap of 100 bp
* 50115 50999: contig of 885 bp in length
* 51000 51099: gap of 100 bp
* 51100 51988: contig of 889 bp in length
* 51989 52088: gap of 100 bp
* 52089 52974: contig of 886 bp in length
* 52975 53074: gap of 100 bp
* 53075 53918: contig of 844 bp in length
* 53919 54018: gap of 100 bp
* 54019 54871: contig of 853 bp in length
* 54872 54971: gap of 100 bp
* 54972 55842: contig of 871 bp in length
* 55843 55942: gap of 100 bp
* 55943 56817: contig of 875 bp in length
* 56818 56917: gap of 100 bp
* 56919 57793: contig of 876 bp in length
* 57794 57893: gap of 100 bp
* 57894 58764: contig of 871 bp in length
* 58765 58864: gap of 100 bp
* 58865 59725: contig of 861 bp in length
* 59726 59825: gap of 100 bp
* 59826 60677: contig of 852 bp in length
* 60678 60777: gap of 100 bp
* 60778 61648: contig of 871 bp in length
* 61649 61748: gap of 100 bp
* 61749 62632: contig of 884 bp in length
* 62633 62732: gap of 100 bp
* 62733 63616: contig of 884 bp in length
* 63617 63716: gap of 100 bp
* 63717 64567: contig of 851 bp in length
* 64568 64667: gap of 100 bp
* 64669 65533: contig of 866 bp in length
* 65534 65633: gap of 100 bp
* 65634 66505: contig of 872 bp in length
* 66506 66605: gap of 100 bp
* 66606 67460: contig of 855 bp in length
* 67461 67560: gap of 100 bp
* 67561 68417: contig of 857 bp in length
* 68418 68517: gap of 100 bp
* 68518 69394: contig of 877 bp in length
* 69395 69494: gap of 100 bp
* 69495 70375: contig of 881 bp in length

Query Match 71.2%; Score 35.6; DB 2; Length 95127;
Best Local Similarity 82.0%; Pred. No. 0.0091;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTTCAGTTCGGGAAACCTGTGCG 50
Db 20592 ATTTAATTCGCTGCGCTCACTGCCCGCTTTCAGTTCGGGAAACCTGTGCG 20641

Search completed: May 10, 2005, 05:52:47
Job time : 306.908 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 01:20:43 ; Search time 54.6776 Seconds
(without alignments)

5413.313 Million cell updates/sec

Title: us-09-896-888a-10

Perfect score: 50

Sequence: 1 acttaagcttatgcgatga.....tccagtcgggaacctgtcg 50

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	50	2	Aaf55402 Sequence
2	36	72.0	16091	5	Aaf55403 Sequence
3	36	72.0	20795	5	Aaf55404 Sequence
4	35.8	71.6	745	3	Aaf55405 Sequence
5	35.8	71.6	791	11	ACN86091
6	35.8	71.6	869	11	ACN83021
7	35.6	71.2	327	4	Aaf55360
8	35.6	71.2	586	5	ADL39252
9	35.6	71.2	1637	5	Aaf55360
10	35.6	71.2	1695	5	Aaf55360
11	35.6	71.2	1942	5	Aaf55360
12	35.6	71.2	2346	5	Aaf55360
13	35.6	71.2	2424	5	Aaf55360
14	35.6	71.2	2710	5	Aaf55360
15	35.6	71.2	2710	5	Aaf55360
16	35.6	71.2	2757	5	Aaf55360
17	35.6	71.2	3859	12	ADL48102
18	35.6	71.2	5909	5	Aaf55360
19	35.6	71.2	10771	5	Aaf55360
20	35.6	71.2	20974	5	Aaf55360

21	35.2	70.4	60	4	Aaf55402	Aaf55402 Sequence
22	35.2	70.4	60	4	Aaf55403	Aaf55403 Sequence
23	35.2	70.4	114	10	ADL18713	ADL18713 Plasmid p
24	35.2	70.4	118	2	AAT01221	AAT01221 Oligonucle
25	35.2	70.4	151	2	AAV62171	AAV62171 HSV-2 str
26	35.2	70.4	173	5	ABV36253	ABV36253 Human pro
27	35.2	70.4	176	4	AAK85632	AAK85632 Human imm
28	35.2	70.4	186	1	AAK85619	AAK85619 Human o
29	35.2	70.4	186	1	AAK85619	AAK85619 Human o
30	35.2	70.4	194	4	AAK85619	AAK85619 Human imm
31	35.2	70.4	195	4	AAK85619	AAK85619 Human dig
32	35.2	70.4	195	4	AAK85619	AAK85619 Human dig
33	35.2	70.4	195	4	AAK85619	AAK85619 Human dig
34	35.2	70.4	195	4	AAK85619	AAK85619 Human imm
35	35.2	70.4	195	4	AAK85619	AAK85619 Human imm
36	35.2	70.4	195	4	AAK85619	AAK85619 Human imm
37	35.2	70.4	195	4	AAK85619	AAK85619 Human imm
38	35.2	70.4	195	4	AAK85619	AAK85619 Human imm
39	35.2	70.4	195	4	AAK85619	AAK85619 Human imm
40	35.2	70.4	195	4	AAK85619	AAK85619 Human imm
41	35.2	70.4	195	4	AAK85619	AAK85619 Human imm
42	35.2	70.4	195	4	AAK85619	AAK85619 Human imm
43	35.2	70.4	195	4	AAK85619	AAK85619 Human imm
44	35.2	70.4	195	4	AAK85619	AAK85619 Human imm
45	35.2	70.4	195	4	AAK85619	AAK85619 Human imm

ALIGNMENTS

RESULT 1
AAV62498
ID AAV62498 standard; DNA; 50 BP.
XX
AC AAV62498;
XX
DT 19-JAN-1999 (first entry)
XX
DE Plasmid p2Zop2J-3 constructing primer 2.
XX
KW Op ie2; promoter; shuttle vector; transformation; melanotransferrin;
KW immediate early baculovirus promoter; prokaryotic; transcription;
KW bleomycin/phleomycin-type antibiotic; insect cell; transposon;
KW ion transport peptide hormone; PCR primer; ss.
XX
OS Synthetic.
XX
PN WO9844141-A2.
XX
PD 08-OCT-1998.
XX
PF 26-MAR-1998; 98WO-CA000282.
XX
PR 27-MAR-1997; 97US-0049946P.
XX
PR 28-JAN-1998; 98CA-02221819.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
XX
PI Grigliatti TA, Theilmann DA, Pfeifer TA, Hegedus DD;
XX
WPI; 1998-557129/47.
XX
DR Expression vectors for transforming insect cells from disparate lines -
XX useful to express heterologous DNA, e.g. to allow study of gene
XX expression and produce commercially important proteins.
XX
PS Disclosure; Page 39; 121pp; English.
XX
CC Primers AAV62497 and AAV62498 were used for the construction of the
XX plasmid p2Zop2J-3. The invention provides a new shuttle vector for
XX transforming insect cells that comprises: (i) prokaryotic origin of
XX replication; (ii) insect promoter having homology to, and capable of
XX functioning as, an immediate early baculovirus promoter; (iii)

CC prokaryotic promoter sequence, and (iv) selectable marker capable of
 CC conferring resistance to a bleomycin/photomycin-type antibiotic under
 CC transcriptional control of (ii) and (iii), in insect and prokaryotic
 CC cells respectively. The vectors can be used to stably transform
 CC (especially insect) cells with heterologous DNA, is useful to allow study
 CC of gene expression and direct expression of heterologous gene products,
 CC such as commercially important proteins. They are especially useful to
 CC allow expression of melanotransferrins, ion transport peptide hormones or
 CC biologically active derivatives in insect cells. They enable
 CC transformation of insect cell lines from disparate species, allowing
 CC screening of lines for optimum post-translational modification of
 CC particular proteins. Shuttle vectors further comprising DNA transposable
 CC elements defining a transposon can be used to optimise heterologous
 CC protein expression and facilitate selection of desired transformants
 XX
 XX Sequence 50 BP; 11 A; 14 C; 12 G; 13 T; 0 U; 0 Other;

Query Match 100.0%; Score 50; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.4e-11;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTAAGCTTATAGCGATGACTGCCGCTTTCCAGTCGGGAAACCTGTGCG 50
 |||||
 DB 1 ACTTAAGCTTATAGCGATGACTGCCGCTTTCCAGTCGGGAAACCTGTGCG 50
 |||||

RESULT 2

AAS89978/c
 ID AAS89978 standard; cDNA; 16091 BP.

XX AC AAS89978;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #25782.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX FN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG25791.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX PS Claim 1; SEQ ID NO 25782; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activities in
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 16091 BP; 4724 A; 3867 C; 3861 G; 3639 T; 0 U; 0 Other;

Query Match 72.0%; Score 36; DB 5; Length 16091;
 Best Local Similarity 88.6%; Pred. No. 0.00011;
 Matches 39; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 GCTTATAGCGATGACTGCCGCTTTCCAGTCGGGAAACCTGTGCG 50
 |||||
 DB 8080 GTTATTGCAATCATTTGCCGCTTTCCAGTCGGGAAACCTGTGCG 8037
 |||||

RESULT 3

AAS92596

ID AAS92596 standard; cDNA; 20795 BP.

XX AC AAS92596;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #28400.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX FN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG28409.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX PS Claim 1; SEQ ID NO 28400; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AA864197-AA894564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 20795 BP; 4255 A; 5208 C; 5716 G; 5614 T; 0 U; 2 Other;
SQ

Query Match 72.0%; Score 36; DB 5; Length 20795;
Best Local Similarity 88.6%; Pred. No. 0.00011;
Matches 39; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 GCTTATGCGATGACGTCGCCGCTTTCAGTCGGGAACCTGTGCG 50
DB 12363 GTTATTGCAATGTCGCCGCTTTCAGTCGGGAACCTGTGCG 12406

RESULT 4

AA87692/c
ID AAA87692 standard; cDNA; 745 BP.

XX
AC AAA87692;

XX 04-DEC-2000 (first entry)

XX Human secreted protein gene 27 SEQ ID NO:37.

XX Human; secreted protein; immunosuppressive; immunostimulant; nootropic;
KW antiinflammatory; cardiant; vulnery; antiulcer; anticonvulsant;
KW antiparkinsonian; neuroprotective; antiviral; antibacterial; cytostatic;
KW antiparasitic; thrombolytic; anticoagulant; antiarteriosclerotic;
KW gene therapy; vaccine; chemotaxis-modulator; angiogenesis-modulator;
KW cancer; immune system disorder; hyperproliferative disorder; infection;
KW cardiovascular disorder; neurological disease; wound healing; ss.

XX Homo sapiens.

XX WO200043495-A2.

XX 27-JUL-2000.

XX 18-JAN-2000; 2000WO-US000903.

XX 19-JAN-1999; 99US-0116330P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Ebner R, Young PE, Ni J, Moore PA;

PI Komatsoulis G, Birse CE;

XX WPI; 2000-499225/44.

DR P-PSDB; AA825691.

XX New isolated polynucleotide encoding a secreted protein useful for
PT preventing, treating or ameliorating a medical condition.

XX Claim 1; Page 394; 451pp; English.

XX The polynucleotide sequences given in AA87666 to AA87708 encodes the
CC human secreted proteins based in AA825665 to AA825755. Human secreted
CC proteins have activities given on the tissues and cells the genes are
CC expressed in. Examples of activities include: immunosuppressive;
CC immunostimulant; antiinflammatory; cardiant; vulnery; antiulcer;
CC nootropic; antiviral; anticonvulsant; antiparkinsonian; neuroprotective;
CC antibacterial; antiparasitic; thrombolytic; anticoagulant;
CC antiarteriosclerotic and cytostatic. The secreted proteins and their
CC polynucleotides can be used in gene therapy and as vaccines, chemotaxis-
CC modulators and angiogenesis- modulators. The human secreted proteins and

CC polynucleotides can be used for diagnosing (the susceptibility to) a
CC pathological condition by determining the presence or absence of a
CC mutation in the polynucleotide or determining the presence or amount of
CC expression of the protein. The polynucleotides and proteins can also be
CC used in the treatment and diagnosis of cancer, diseases of the immune
CC system, hyperproliferative disorders, cardiovascular disorders and
CC neurological disease. They can also be used to promote wound healing and
CC to fight infection. AA87657 to AA87665 and AA825664 represent sequences
CC used in the exemplification of the present invention

SQ Sequence 745 BP; 191 A; 214 C; 188 G; 142 T; 0 U; 10 Other;

Query Match 71.6%; Score 35.8; DB 3; Length 745;
Best Local Similarity 83.3%; Pred. No. 5.9e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 TTAAGCTTATACGATGACTGCCGCTTTCAGTCGGGAACCTGTGCG 50
DB 168 TTAATNGCGTTCGCTCACTGCCGCTTTCAGTCGGGAACCTGTGCG 121

RESULT 5

ACN86091
ID ACN86091 standard; DNA; 791 BP.

XX
AC ACN86091;

XX 02-DEC-2004 (first entry)

XX Breast cancer related marker, seq id 7241.

XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.

XX Homo sapiens.

XX US2003099974-A1.

XX 29-MAY-2003.

XX 18-JUL-2002; 2002US-00198846.

XX 18-JUL-2001; 2001US-0306220P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2003-787014/74.

XX Novel isolated polypeptide associated with breast cancer, useful for
PT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.

XX Disclosure; SEQ ID NO 7241; 36pp; English.

XX The invention relates to an isolated polypeptide (I) associated with
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (I) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN78851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?DocId=2003009974

SQ Sequence 791 BP; 188 A; 198 C; 211 G; 182 T; 0 U; 12 Other;

Query Match 71.6%; Score 35.8; DB 11; Length 791;
Best Local Similarity 83.3%; Pred. No. 6e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
QY 3 TTAAGCTTATAGCGATGACTGCCGCTTTCCAGTCGGGAAACCTGTGCG 50
|||||
DB 544 TTAATTCGCTGCGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTGCG 591

RESULT 6
ACN83021
ID ACN83021 standard; DNA; 869 BP.
XX
AC ACN83021;
XX
XX 02-DEC-2004 (first entry)
XX
XX Breast cancer related marker, seq id 4171.
XX
XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX
XX Homo sapiens.
XX
XX US2003099974-A1.
XX
XX 29-MAY-2003.
XX
XX 18-JUL-2002; 2002US-00198846.
XX
XX 18-JUL-2001; 2001US-0306220P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2003-787014/74.
XX
XX Novel isolated polypeptide associated with breast cancer, useful for
XX detecting presence of polypeptide in sample, as a marker for breast
XX cancer.
XX
XX Disclosure; SEQ ID NO 4171; 36pp; English.
XX
XX The invention relates to an isolated polypeptide (I) associated with
XX breast cancer which is encoded by a nucleic acid molecule comprising a
XX nucleotide sequence (S1). Further disclosed is an antibody that binds to
XX the polypeptide of the invention. The activity of the polypeptide of the
XX invention may be described as cytostatic. The antibody is useful for
XX detecting the presence of (I) in a sample. Nucleic acid molecules of the
XX invention are useful in the detection of breast tumours. (I) is useful as
XX a marker for breast cancer and in breast cancer therapy. Sequences given
XX in records ACN78851-ACN92934 represent nucleic acid markers associated
XX with breast cancer. Note: The sequence listing does not form part of the
XX specification but may be obtained in electronic format from the USPTO web
XX site at seqdata.uspto.gov/sequence.html?docID=20030099974
XX
XX Sequence 869 BP; 206 A; 202 C; 225 G; 216 T; 0 U; 20 Other;
XX
XX Query Match 71.6%; Score 35.8; DB 11; Length 869;
XX Best Local Similarity 83.3%; Pred. No. 6.1e-05;
XX Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
XX
QY 3 TTAAGCTTATAGCGATGACTGCCGCTTTCCAGTCGGGAAACCTGTGCG 50
|||||
DB 569 TTAATTCGCTGCGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTGCG 616

RESULT 7
AAF55360
ID AAF55360 standard; DNA; 327 BP.
XX
AC AAF55360;
XX
XX 29-MAY-2001 (first entry)
XX
XX Sequence obtained after sequencing with modified Taq DNA polymerase.
```

```
XX Thermostable DNA polymerase; signal uniformity; salt tolerance;
XX nucleic acid amplification; polymerase chain reaction; pREFV2pref;
XX self-sustained sequence replication; DNA sequencing; 88.
XX
XX Unidentified.
XX
XX WO200114568-A1.
XX
XX 01-MAR-2001.
XX
XX 10-AUG-2000; 2000WO-US022150.
XX
XX 21-AUG-1999; 99US-0150167P.
XX
XX 17-SEP-1999; 99US-0154739P.
XX
XX (AMSH) AMERSHAM PHARMACIA BIOTECH INC.
XX
XX Davis M, Nelson J, Kumar S, Finn PJ, Nampalli S, Flicke P;
XX WPI; 2001-226620/23.
XX
XX New purified recombinant thermostable DNA polymerases having amino acid
XX substitutions at E410R or E681R, useful in recombinant DNA techniques,
XX e.g. nucleic acid amplification or high temperature DNA sequencing.
XX
XX Example 1; Fig 7; 48pp; English.
XX
XX The specification describes new purified recombinant thermostable DNA
XX polymerases having an amino acid substitution at E681. The new DNA
XX polymerases have improved discrimination properties (and thus resulting
XX in improved signal uniformity) and increased tolerance to high salt
XX conditions. They also modulate the incorporation of terminators having a
XX net positive or net negative charge during sequencing. The recombinant
XX thermostable DNA polymerases are useful in many recombinant DNA
XX techniques, e.g. nucleic acid amplification by polymerase chain reaction,
XX self-sustained sequence replication, or high temperature DNA sequencing.
XX The recombinant thermostable DNA polymerases are also useful in
XX increasing the uniformity of dye-terminator incorporation in fluorescent
XX dye DNA sequencing. The present sequence represents a sequence obtained
XX after sequencing with a modified Taq DNA polymerase of the invention
XX
XX Sequence 327 BP; 75 A; 82 C; 89 G; 81 T; 0 U; 0 Other;
XX
XX Query Match 71.2%; Score 35.6; DB 4; Length 327;
XX Best Local Similarity 82.0%; Pred. No. 5.8e-05;
XX Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
XX
QY 1 ACTTAAGCTTATAGCGATGACTGCCGCTTTCCAGTCGGGAAACCTGTGCG 50
|||||
DB 127 ATTTAATTCGCTGCGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTGCG 176

RESULT 8
ADL39252
ID ADL39252 standard; DNA; 586 BP.
XX
AC ADL39252;
XX
XX 20-MAY-2004 (first entry)
XX
XX Human ovarian cancer DNA marker #13142.
XX
XX Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX
XX Homo sapiens.
XX
XX WO200170979-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US009126.
XX
```

PR 21-MAR-2000; 2000US-0191031P.
PR 25-MAY-2000; 2000US-0207124P.
PR 15-JUN-2000; 2000US-0211940P.
PR 07-JUL-2000; 2000US-0216820P.
PR 25-JUL-2000; 2000US-0220661P.
PR 21-DEC-2000; 2000US-0257672P.
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lee J, Lillie J;
XX
XX WPI; 2001-611502/70.
XX
PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT cancer cells as compared to their normal non-cancerous ovarian cells are
PT used to characterize stage, grade, histological type of ovarian cancer.
XX
XX Disclosure; SEQ ID NO 13142; 106pp; English.
XX
XX The invention relates to nucleic acid markers which are overexpressed in
XX ovarian cancer cells as compared to their expression in normal (i.e. non-
XX cancerous) ovarian cells. The invention also relates to polypeptides
XX encoded by the markers, antibodies that selectively bind to the
XX polypeptides, a method of inhibiting ovarian cancer in a patient at risk
XX of developing ovarian cancer involving inhibiting expression of a gene
XX corresponding to a marker of the invention and a method of treating a
XX patient afflicted with ovarian cancer comprising providing to cells of
XX the patient an antisense oligonucleotide complementary to a marker of the
XX invention. The markers are useful for assessing if a patient is afflicted
XX with ovarian cancer, which involves comparing the level of expression of
XX a marker in a patient sample and a normal level of expression of the
XX expression levels indicates ovarian cancer. A difference between the
XX marker corresponds to a secreted protein or to a transcribed
XX polynucleotide or its portion. The level of expression of the marker is
XX assessed by detecting the presence in the sample, a protein or protein
XX fragment corresponding to the marker. The presence of protein or protein
XX fragment is detected using an antibody that specifically binds with the
XX protein or protein fragment. Alternatively, the level of expression of
XX the marker is assessed by detecting the presence of a transcribed
XX polynucleotide which anneals with the marker or anneals with a portion of
XX the polynucleotide comprising the marker, under stringent conditions. The
XX marker is also used for monitoring the progression of ovarian cancer in a
XX patient which involves detecting expression of the marker in a patient
XX sample at a first point in time, repeating the method at a subsequent
XX time and comparing the level of expression. The method is carried out
XX using an ovarian tissue sample. A composition comprising a marker,
XX polypeptide or antibody of the invention is used to treat ovarian cancer.
XX This sequence represents a human ovarian cancer DNA marker of the
XX invention.
SQ Sequence 586 BP; 168 A; 131 C; 141 G; 146 T; 0 U; 0 Other;
Query Match 71.2%; Score 35.6; DB 5; Length 586;
Best Local Similarity 82.0%; Pred. No. 6.7e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ACTTAGCTTATAGGATGACCTGCGCCGCTTTCAGTCCGGAAACCTGTGCG 50
DB 516 ATTAAATCGGTGGCTCACTGCCCGCTTTCAGTCCGGAAACCTGTGCG 565
RESULT 9
AAS77556/c
ID AAS77556 standard; cDNA; 1637 BP.
XX
XX AAS77556;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #13360.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
OS
XX WO200175067-A2.
PN
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-US008631.
PF
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
PR
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ABG13369.
DR
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
PS Claim 1; SEQ ID NO 13360; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 1637 BP; 447 A; 437 C; 395 G; 357 T; 0 U; 1 Other;
Query Match 71.2%; Score 35.6; DB 5; Length 1637;
Best Local Similarity 82.0%; Pred. No. 8.7e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ACTTAGCTTATAGGATGACCTGCGCCGCTTTCAGTCCGGAAACCTGTGCG 50
DB 359 ATTAAATCGGTGGCTCACTGCCCGCTTTCAGTCCGGAAACCTGTGCG 310
RESULT 10
AAS77545/c
ID AAS77545 standard; cDNA; 1695 BP.
XX
XX AAS77545;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #13349.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX

```
OS Homo sapiens.
XX WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Dmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG13358.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 1; SEQ ID NO 13349; 103pp; English.
XX PS The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX CC coding sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1695 BP; 415 A; 463 C; 426 G; 390 T; 0 U; 1 Other;
XX Query Match 71.2%; Score 35.6; DB 5; Length 1695;
XX Best Local Similarity 82.0%; Pred. No. 8.8e-05;
XX Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
XX QY 1 ACTTAAGCTTATAGCGATGACTGCCGCTTTCAGTCGGGAACCTGTGC 50
XX Db 651 AATTAATTGCGTTGGCTCACTGCCGCTTTCAGTCGGGAACCTGTGC 602
XX RESULT 11
XX AAS69138/c
XX ID AAS69138 standard; cDNA; 1942 BP.
XX AC AAS69138;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #4942.
XX DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX XX 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Dmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG04951.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 1; SEQ ID NO 4942; 103pp; English.
XX PS The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX CC coding sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1942 BP; 516 A; 524 C; 495 G; 407 T; 0 U; 0 Other;
XX Query Match 71.2%; Score 35.6; DB 5; Length 1942;
XX Best Local Similarity 82.0%; Pred. No. 9.1e-05;
XX Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
XX QY 1 ACTTAAGCTTATAGCGATGACTGCCGCTTTCAGTCGGGAACCTGTGC 50
XX Db 1193 AATTAATTGCGTTGGCTCACTGCCGCTTTCAGTCGGGAACCTGTGC 1134
XX RESULT 12
XX AAS77547
XX ID AAS77547 standard; cDNA; 2346 BP.
XX AC AAS77547;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #13351.
XX DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX XX
```

```

PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
PR (HYSE-) HYSEQ INC.
XX
PA Drmanac RT, Liu C, Tang YT;
XX
PI WPI; 2001-639362/73.
XX
PI P-PSDB; ABG13360.
XX
DR New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
PS Claim 1; SEQ ID NO 13351; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2346 BP; 542 A; 590 C; 609 G; 605 T; 0 U; 0 Other;

Query Match 71.2%; Score 35.6; DB 5; Length 2346;
Best Local Similarity 82.0%; Pred. No. 9.6e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ACTTAAGCTTATAGCGATGACTCCCGCTTTCAGTCGGGAACCTGTGCG 50
Db 1672 AATTAAATGCGTGGCTCACTCCCGCTTTCAGTCGGGAACCTGTGCG 1721

RESULT 13
AAS87523/c
ID AAS87523 standard; cDNA; 2424 BP.
XX
AC AAS87523;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #23327.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX

```

```

PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
PR (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
PI WPI; 2001-639362/73.
XX
PI P-PSDB; ABG23336.
XX
DR New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
PS Claim 1; SEQ ID NO 23327; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2424 BP; 528 A; 721 C; 738 G; 437 T; 0 U; 0 Other;

Query Match 71.2%; Score 35.6; DB 5; Length 2424;
Best Local Similarity 82.0%; Pred. No. 9.7e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ACTTAAGCTTATAGCGATGACTCCCGCTTTCAGTCGGGAACCTGTGCG 50
Db 1365 AATTAAATGCGTGGCTCACTCCCGCTTTCAGTCGGGAACCTGTGCG 1316

RESULT 14
AAS77561/c
ID AAS77561 standard; cDNA; 2710 BP.
XX
AC AAS77561;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #13365.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX

```

```
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX
DR P-PSDB; ABG13374.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 13365; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AA864197-AA894564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2710 BP; 615 A; 724 C; 690 G; 681 T; 0 U; 0 Other;
Query Match 71.2%; Score 35.6; DB 5; Length 2710;
Best Local Similarity 82.0%; Pred. No. 9.9e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 50
Db 1080 AATTAATGCGTTGGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 1031
RESULT 15
ADE09733/c
ID ADE09733 standard; DNA; 2710 BP.
XX
AC ADE09733;
XX
XX
DT 29-JAN-2004 (first entry)
XX
DE Novel DNA-related contig nucleotide sequence #455.
XX
KW novel gene; novel protein; tissue marker; molecular weight marker;
KW chromosome marker; genetic disorder; contig; ds.
XX
OS Unidentified.
XX
FN WO20003054152-A2.
XX
XX 03-JUL-2003.
XX
XX 10-DEC-2002; 2002WO-US039555.
XX
XX 10-DEC-2001; 2001US-0339739P.
PR 11-DEC-2001; 2001US-0339453P.
PR
```

```
PR 14-MAR-2002; 2002US-0365091P.
PR 14-MAR-2002; 2002US-0365384P.
PR 12-APR-2002; 2002US-0372381P.
PR 12-APR-2002; 2002US-0372615P.
PR 22-APR-2002; 2002US-00128558.
PR 24-APR-2002; 2002US-0376045P.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
XX Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
XX Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX WPI; 2003-569235/53.
XX
XX New polynucleotides, useful for expressing recombinant proteins for
XX analysis, characterization or therapeutic use, or as markers for tissues
XX in which the corresponding protein is preferentially expressed.
XX
XX Disclosure; SEQ ID NO 2277; 1177pp; English.
XX
XX The invention comprises the amino acid and coding sequences of novel
XX proteins. The DNA and protein sequences of the invention are useful as:
XX markers for tissues in which the corresponding protein is preferentially
XX expressed; as molecular weight markers on gels; as chromosome markers or
XX tags; to identify chromosomes or to map related gene positions; and to
XX compare with endogenous DNA sequences in patients to identify potential
XX genetic disorders. The present DNA sequence was used in the
XX exemplification of the invention.
XX
XX Sequence 2710 BP; 615 A; 724 C; 690 G; 681 T; 0 U; 0 Other;
Query Match 71.2%; Score 35.6; DB 10; Length 2710;
Best Local Similarity 82.0%; Pred. No. 9.9e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 50
Db 1080 AATTAATGCGTTGGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 1031
Search completed: May 10, 2005, 04:32:55
Job time : 61.6776 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 04:19:16 ; Search time 16.6245 Seconds
(without alignments)
4921.273 Million cell updates/sec

Title: US-09-896-888A-10
Perfect score: 50
Sequence: 1 acttaagttatagcgatga.....tccagtcgggaacctgtcg 50

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCUS COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	35.2	70.4	114	2	US-08-778-217-1
C 2	35.2	70.4	114	2	US-08-821-948-1
C 3	35.2	70.4	114	2	US-09-397-955C-1
C 4	35.2	70.4	118	5	PCT-US95-04092-1
C 5	35.2	70.4	201	1	US-08-276-852-41
C 6	35.2	70.4	201	1	US-08-133-011-114
C 7	35.2	70.4	201	1	US-08-322-730A-114
C 8	35.2	70.4	201	1	US-08-387-874-87
C 9	35.2	70.4	201	1	US-08-899-575-41
C 10	35.2	70.4	201	1	US-08-899-575-41
C 11	35.2	70.4	201	2	US-08-383-619-114
C 12	35.2	70.4	201	3	US-08-907-739-114
C 13	35.2	70.4	201	3	US-09-729-597-114
C 14	35.2	70.4	201	5	PCT-US93-08364-87
C 15	35.2	70.4	201	5	PCT-US95-08743-41
C 16	35.2	70.4	221	1	US-08-531-132-1
C 17	35.2	70.4	304	4	US-08-120-324-15
C 18	35.2	70.4	310	4	US-10-165-857-1
C 19	35.2	70.4	310	4	US-10-165-856A-1
C 20	35.2	70.4	352	4	US-10-165-857-2
C 21	35.2	70.4	352	4	US-10-165-856A-2
C 22	35.2	70.4	357	3	US-09-525-046-3
C 23	35.2	70.4	360	2	US-08-182-173A-1
C 24	35.2	70.4	450	4	US-09-486-336A-2
C 25	35.2	70.4	504	2	US-08-768-550-12
C 26	35.2	70.4	505	2	US-08-768-550-11
C 27	35.2	70.4	506	2	US-08-768-550-10

28	35.2	70.4	584	4	US-09-702-705-639	Sequence 639, App
29	35.2	70.4	584	4	US-09-736-457-639	Sequence 639, App
30	35.2	70.4	584	4	US-09-614-124B-639	Sequence 639, App
31	35.2	70.4	584	4	US-09-671-325-639	Sequence 639, App
32	35.2	70.4	584	4	US-09-589-184-639	Sequence 639, App
33	35.2	70.4	584	4	US-09-658-824-639	Sequence 639, App
34	35.2	70.4	585	4	US-09-334-818A-7	Sequence 7, Appli
35	35.2	70.4	591	4	US-09-334-818A-3	Sequence 3, Appli
36	35.2	70.4	597	4	US-09-334-818A-10	Sequence 10, Appl
37	35.2	70.4	597	4	US-09-334-818A-16	Sequence 16, Appl
38	35.2	70.4	598	4	US-09-334-818A-11	Sequence 11, Appl
39	35.2	70.4	599	4	US-09-334-818A-6	Sequence 6, Appli
40	35.2	70.4	601	4	US-09-334-818A-18	Sequence 18, Appl
41	35.2	70.4	601	4	US-09-334-818A-19	Sequence 19, Appl
42	35.2	70.4	602	4	US-09-334-818A-14	Sequence 14, Appl
43	35.2	70.4	602	4	US-09-334-818A-15	Sequence 15, Appl
44	35.2	70.4	602	4	US-09-334-818A-22	Sequence 22, Appl
45	35.2	70.4	604	4	US-09-334-818A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-778-217-1/c
; Sequence 1, Application US/08778217
; Patent No. 5935833
; GENERAL INFORMATION:
; APPLICANT: Kacian et al.
; TITLE OF INVENTION: Highly-Purified Recombinant
; TITLE OF INVENTION: Reverse Transcriptase
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gen-Probe Incorporated
; STREET: 10210 Genetic Center Drive
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92121-4362
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Microsoft MS-DOS (Version 6.0)
; SOFTWARE: FASTSEQ
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/778,217
; FILING DATE: January 9, 1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/221,804
; FILING DATE: April 1, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Christine A. Gritzmacher
; REGISTRATION NUMBER: 40,627
; REFERENCE/DOCKET NUMBER: MOL2A-A01F01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 410-8926
; TELEFAX: (619) 410-8928
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-778-217-1

Query Match 70.4%; Score 35.2; DB 2; Length 114;
Best Local Similarity 83.3%; Pred. No. 9.6e-06;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 3 TTAAGCTTATAGCGATGACTGCCCGTTTCCAGTCGGAAACCTGTGCG 50

Db 63 TTAATTCGTTGCGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTGCG 16

RESULT 2

US-08-821-948-1/c

; Sequence 1, Application US/08821948

; Patent No. 5998195

; GENERAL INFORMATION:

; APPLICANT: Kacian et al.

; TITLE OF INVENTION: Highly-Purified Recombinant

; TITLE OF INVENTION: Reverse Transcriptase

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Gen-Probe Incorporated

; STREET: 10210 Genetic Center Drive

; CITY: San Diego

; STATE: California

; COUNTRY: USA

; ZIP: 92121-4362

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

; COMPUTER: IBM compatible

; OPERATING SYSTEM: Microsoft MS-DOS (Version 6.0)

; SOFTWARE:

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/821,948

; FILING DATE: March 22, 1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA: including application

; PRIOR APPLICATION DATA: described below:

; APPLICATION NUMBER: 08/443,781

; FILING DATE: May 18, 1995

; APPLICATION NUMBER: 08/221,804

; FILING DATE: April 1, 1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Christine A. Gritzmacher

; REGISTRATION NUMBER: 40,627

; REFERENCE/DOCKET NUMBER: M012A (New Ref.: GP059-04.FW2)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 410-8926

; TELEFAX: (619) 410-8928

; INFORMATION FOR SEQ ID NO: 1:

; LENGTH: 114

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-821-948-1

Query Match 70.4%; Score 35.2; DB 2; Length 114;
Best Local Similarity 83.3%; Pred. No. 9.6e-06;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGCGATGACCTCCCGCTTTCCAGTCGGGAAACCTGTGCG 50

Db 63 TTAATTCGTTGCGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTGCG 16

RESULT 3

US-09-397-955C-1/c

; Sequence 1, Application US/09397955C

; Patent No. 6593120

; GENERAL INFORMATION:

; APPLICANT: RIGGS, Michael G.

; APPLICANT: SORESENSEN, Matthew

; TITLE OF INVENTION: RECOMBINANT DNA ENCODING REVERSE TRANSCRIPTASE DERIVED FROM

; TITLE OF INVENTION: MOLONEY MURINE LEUKEMIA VIRUS

; FILE REFERENCE: GP059-05.CP1

; CURRENT APPLICATION NUMBER: US/09/397,955C

; CURRENT FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: 08/821,948

; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: 08/443,781
; PRIOR FILING DATE: 1995-05-18
; PRIOR APPLICATION NUMBER: 08/221,804
; PRIOR FILING DATE: 1994-04-01
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 114
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Oligonucleotide used to construct plasmid pUC 18N
US-09-397-955C-1

Query Match 70.4%; Score 35.2; DB 4; Length 114;
Best Local Similarity 83.3%; Pred. No. 9.6e-06;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGCGATGACCTCCCGCTTTCCAGTCGGGAAACCTGTGCG 50

Db 63 TTAATTCGTTGCGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTGCG 16

RESULT 4

PCT-US95-04092-1/c

; Sequence 1, Application PC/TUS9504092

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: HIGHLY-PURIFIED RECOMBINANT REVERSE

; TITLE OF INVENTION: TRANSCRIPTASE

; NUMBER OF SEQUENCES: 18

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; SOFTWARE:

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/04092

; FILING DATE:

; CLASSIFICATION:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 118 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; PCT-US95-04092-1

Query Match 70.4%; Score 35.2; DB 5; Length 118;
Best Local Similarity 83.3%; Pred. No. 9.7e-06;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGCGATGACCTCCCGCTTTCCAGTCGGGAAACCTGTGCG 50

Db 63 TTAATTCGTTGCGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTGCG 16

RESULT 5

US-08-276-852-41/c

; Sequence 41, Application US/08276852

; Patent No. 5652138

; GENERAL INFORMATION:

; APPLICANT: Burton, Dennis R

; APPLICANT: Barbas, Carlos F

; APPLICANT: Lerner, Richard A

; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES

; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS

; NUMBER OF SEQUENCES: 170

; CORRESPONDENCE ADDRESS:

ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,852
FILING DATE: 18-JUL-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCR1452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-276-852-41

Query Match 70.4%; Score 35.2; DB 1; Length 201;
Best Local Similarity 83.3%; Pred. No. 1.1e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 3 TTAAGCTTATAGCGACTGCCCGCTTCCAGTCGGGAAACCTGTCG 50
DB 65 TTAATTGCGTTGCGCTCACTGCCCGCTTCCAGTCGGGAAACCTGTCG 18

RESULT 6
US-08-133-011-114/c
Sequence 114, Application US/08133011
Patent No. 5658727
GENERAL INFORMATION:
APPLICANT: Kang, Angray
APPLICANT: Barbas, Carlos
APPLICANT: Lerner, Richard A.
TITLE OF INVENTION: HETERODIMERIC RECEPTOR LIBRARIES USING
NUMBER OF SEQUENCES: 161
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10666 No. 5658727th Torrey Pines Road, TPC-8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,852
FILING DATE: 18-JUL-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCR1452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-276-852-41

SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/133,011
FILING DATE: 08-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/683,602
FILING DATE: 10-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/826,623
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US 92/03091
FILING DATE: 10-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRF 238.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 114:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-133-011-114

Query Match 70.4%; Score 35.2; DB 1; Length 201;
Best Local Similarity 83.3%; Pred. No. 1.1e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 3 TTAAGCTTATAGCGACTGCCCGCTTCCAGTCGGGAAACCTGTCG 50
DB 65 TTAATTGCGTTGCGCTCACTGCCCGCTTCCAGTCGGGAAACCTGTCG 18

RESULT 7
US-08-322-730A-114/c
Sequence 114, Application US/08322730A
Patent No. 5759817
GENERAL INFORMATION:
APPLICANT: Barbas, Carlos
TITLE OF INVENTION: HETERODIMERIC RECEPTOR LIBRARIES USING
NUMBER OF SEQUENCES: 125
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10550 No. 5759817th Torrey Pines Road, Suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322,730A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/826,623
FILING DATE: 27-JAN-1992
APPLICATION NUMBER: US 07/683,602
FILING DATE: 10-APR-1991

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Bingham, Douglas A
; REGISTRATION NUMBER: 32,457
; REFERENCE/DOCKET NUMBER: SC0707P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-784-2937
; TELEFAX: 619-784-9399
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; US-08-322-730A-114

Query Match          70.4%; Score 35.2; DB 1; Length 201;
Best Local Similarity 83.3%; Pred. No. 1.1e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 3 TTAAGCTTATAGCATGACTCCCGCTTCCAGTCGGGAAACCTGTGCG 50
Db 65 TTAATTGCGTTCGCTCACTCCCGCTTCCAGTCGGGAAACCTGTGCG 18

RESULT 8
US-08-387-874-87/c
; Sequence 87, Application US/08387874
; Patent No. 5770356
; GENERAL INFORMATION:
; APPLICANT: Light, Paul L., II
; APPLICANT: Lerner, Richard A.
; TITLE OF INVENTION: PHAGMIDS COEXPRESSING A SURFACE
; TITLE OF INVENTION: RECEPTOR AND A SURFACE HETEROLOGOUS PROTEIN
; NUMBER OF SEQUENCES: 97
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 NO. 5770356th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/387,874
; FILING DATE: 22-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08364
; FILING DATE: 03-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/941,369
; FILING DATE: 04-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 303.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 87:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEtical: NO
; ANTI-SENSE: NO
; US-08-387-874-87

Query Match          70.4%; Score 35.2; DB 1; Length 201;
Best Local Similarity 83.3%; Pred. No. 1.1e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 3 TTAAGCTTATAGCATGACTCCCGCTTCCAGTCGGGAAACCTGTGCG 50
Db 65 TTAATTGCGTTCGCTCACTCCCGCTTCCAGTCGGGAAACCTGTGCG 18

RESULT 9
US-08-899-575-41/c
; Sequence 41, Application US/08899575
; Patent No. 5770440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 NO. 5770440th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEtical: NO
; ANTI-SENSE: NO
; US-08-899-575-41

Query Match          70.4%; Score 35.2; DB 1; Length 201;
Best Local Similarity 83.3%; Pred. No. 1.1e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

QY 3 TTAAGCTTATAGCGATGACGCGCTTCCAGTCGGGAACCTGTGC 50
Db 65 TTAATTGCGTGGCTCACTGCGCGCTTCCAGTCGGGAACCTGTGC 18

RESULT 10

US-08-899-575-41/c
; Sequence 41, Application US/08899575
; Patent No. 5804440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5804440th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION/DOCKET NUMBER: SCRI452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-899-575-41

Query Match 70.4%; Score 35.2; DB 1; Length 201;
Best Local Similarity 83.3%; Pred. No. 1.1e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGCGATGACGCGCTTCCAGTCGGGAACCTGTGC 50
Db 65 TTAATTGCGTGGCTCACTGCGCGCTTCCAGTCGGGAACCTGTGC 18

RESULT 11

US-08-383-619-114/c
; Sequence 114, Application US/08383619
; Patent No. 5955341

; GENERAL INFORMATION:
; APPLICANT: Kang, Angray
; APPLICANT: Barbas, Carlos
; APPLICANT: Lerner, Richard
; TITLE OF INVENTION: HETERODIMERIC RECEPTOR LIBRARIES USING
; TITLE OF INVENTION: PHAGEMIDS
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DOUGLAS A. BINGHAM
; STREET: 11300 Sorrento Valley Road, Suite 200
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92121

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/383,619
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/126,680
; FILING DATE:
; APPLICATION NUMBER: US/07/683,602
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bingham, Douglas A.
; REGISTRATION/DOCKET NUMBER: 32,457
; REFERENCE/DOCKET NUMBER: SCR0371P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-1555

INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-383-619-114

Query Match 70.4%; Score 35.2; DB 2; Length 201;
Best Local Similarity 83.3%; Pred. No. 1.1e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGCGATGACGCGCTTCCAGTCGGGAACCTGTGC 50
Db 65 TTAATTGCGTGGCTCACTGCGCGCTTCCAGTCGGGAACCTGTGC 18

RESULT 12

US-08-907-739-114/c
; Sequence 114, Application US/08907739
; Patent No. 6235469
; GENERAL INFORMATION:
; APPLICANT: Kang, Angray
; APPLICANT: Barbas, Carlos
; APPLICANT: Lerner, Richard A.
; TITLE OF INVENTION: HETERODIMERIC RECEPTOR LIBRARIES USING
; TITLE OF INVENTION: PHAGEMIDS
; NUMBER OF SEQUENCES: 161
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 6235469th Torrey Pines Road, TPC-8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/907,739
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/133,011
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/826,623
; FILING DATE: 27-JAN-1992
; APPLICATION DATA:
; APPLICATION NUMBER: PCT/US 92/03091
; FILING DATE: 10-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRF 238.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-907-739-114

Query Match 70.4%; Score 35.2; DB 3; Length 201;
Best Local Similarity 83.3%; Pred. No. 1.1e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGCATGCTCCCGCTTTCAGTCGGGAAACCTGTGC 50
DB 65 TTAATTCGGTTCCTCACTCCCGCTTTCAGTCGGGAAACCTGTGC 18

RESULT 13
US-09-729-597-114/c
; Sequence 114, Application US/09729597
; Patent No. 6468738
; GENERAL INFORMATION:
; APPLICANT: Kang, Angray
; Barbas, Carlos
; Lerner, Richard A.
; TITLE OF INVENTION: HETERODIMERIC RECEPTOR LIBRARIES USING
; PHAGEMIDS
; NUMBER OF SEQUENCES: 161
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; Patent Counsel
; STREET: 10666 No. 6468738th Torrey Pines Road, TPC-8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/729,597
; FILING DATE: 04-Dec-2000
; CLASSIFICATION: 435
```

```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/133,011
; FILING DATE: 1994-09-29
; APPLICATION NUMBER: US 07/826,623
; FILING DATE: 27-JAN-1992
; APPLICATION NUMBER: PCT/US 92/03091
; FILING DATE: 10-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRF 238.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 114:
US-09-729-597-114

Query Match 70.4%; Score 35.2; DB 3; Length 201;
Best Local Similarity 83.3%; Pred. No. 1.1e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGCATGCTCCCGCTTTCAGTCGGGAAACCTGTGC 50
DB 65 TTAATTCGGTTCCTCACTCCCGCTTTCAGTCGGGAAACCTGTGC 18

RESULT 14
PCT-US93-08364-87/c
; Sequence 87, Application PC/TUS9308364
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PHAGEMIDS COEXPRESSING A SURFACE
; RECEPTOR AND A SURFACE HETEROLOGOUS PROTEIN
; NUMBER OF SEQUENCES: 97
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08364
; FILING DATE: 03-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/941,369
; FILING DATE: 04-SEP-1992
; INFORMATION FOR SEQ ID NO: 87:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; PCT-US93-08364-87

Query Match 70.4%; Score 35.2; DB 5; Length 201;
Best Local Similarity 83.3%; Pred. No. 1.1e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGCATGCTCCCGCTTTCAGTCGGGAAACCTGTGC 50
DB 65 TTAATTCGGTTCCTCACTCCCGCTTTCAGTCGGGAAACCTGTGC 18
```

RESULT 15
PCT-US95-08743-41/c
; Sequence 41, Application PC/TUS9508743
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08743
; FILING DATE: 11-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
PCT-US95-08743-41

Query Match 70.4%; Score 35.2; DB 5; Length 201;
Best Local Similarity 83.3%; Pred. No. 1.1e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 3 TTAAAGCTTATAGCGATGACTGCCCGCTTCCAGTCGGGAAACCTGTGCG 50
Db 65 TTAAATTCGCTTGCCTCACTGCCCGCTTCCAGTCGGGAAACCTGTGCG 18

Search completed: May 10, 2005, 07:29:36
Job time : 17.6245 secs

2

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 05:52:55 ; Search time 141.277 Seconds
(without alignments)
2164.037 Million cell updates/sec

Title: US-09-896-888A-10
Perfect score: 50
Sequence: 1 acttaagcttatagcatga.....tccagtcgggaacctgtcg 50

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 5654200 seqs, 3057283753 residues

Total number of hits satisfying chosen parameters: 11308400

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications_NA:*
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
 - 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
 - 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
 - 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
 - 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
 - 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
 - 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
 - 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
 - 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
 - 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
 - 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
 - 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
 - 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
 - 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	50	9	US-09-896-888A-10
2	35.8	71.6	745	11	US-09-896-888A-10
3	35.8	71.6	791	14	US-09-896-888A-10
4	35.8	71.6	869	14	US-10-198-846-7241
5	35.6	71.2	415	18	US-10-198-846-4171
6	35.6	71.2	586	10	US-10-437-963-80752
7	35.6	71.2	865	13	US-09-814-353-13142
8	35.6	71.2	865	13	US-10-027-632-2086
9	35.6	71.2	865	13	US-10-027-632-2087
10	35.6	71.2	865	17	US-10-027-632-2086
11	35.6	71.2	865	17	US-10-027-632-2087
					Sequence 8806, A

C	12	35.6	71.2	2721	18	US-10-437-963-39259	Sequence 39259, A
	13	35.2	70.4	173	18	US-10-357-930-36271	Sequence 36271, A
	14	35.2	70.4	179	18	US-10-357-930-45266	Sequence 45266, A
	15	35.2	70.4	195	9	US-09-764-846-347	Sequence 347, App
	16	35.2	70.4	195	10	US-09-764-872-710	Sequence 710, App
	17	35.2	70.4	195	10	US-09-764-891-7492	Sequence 7492, App
	18	35.2	70.4	195	10	US-09-764-891-9818	Sequence 9818, App
	19	35.2	70.4	195	14	US-10-091-483-347	Sequence 347, App
	20	35.2	70.4	195	15	US-10-205-428-959	Sequence 959, App
	21	35.2	70.4	201	16	US-10-016-986-41	Sequence 41, Appl
	22	35.2	70.4	201	17	US-10-273-973-114	Sequence 114, App
	23	35.2	70.4	203	18	US-10-357-930-36213	Sequence 36213, A
	24	35.2	70.4	205	18	US-10-333-872A-286	Sequence 286, App
	25	35.2	70.4	213	9	US-09-969-617-1	Sequence 1, Appl1
	26	35.2	70.4	238	17	US-10-319-227A-8	Sequence 8, Appl1
	27	35.2	70.4	238	17	US-10-319-227A-44	Sequence 44, Appl
	28	35.2	70.4	238	17	US-10-286-549A-8	Sequence 8, Appl1
	29	35.2	70.4	238	17	US-10-286-549A-44	Sequence 44, Appl
	30	35.2	70.4	244	18	US-10-357-930-45545	Sequence 45545, A
	31	35.2	70.4	253	18	US-10-357-930-38331	Sequence 38331, A
	32	35.2	70.4	265	9	US-09-924-035A-642	Sequence 642, App
	33	35.2	70.4	271	17	US-10-653-677-7	Sequence 7, Appl
	34	35.2	70.4	286	9	US-09-924-035A-66	Sequence 66, Appl
	35	35.2	70.4	297	18	US-10-437-963-21902	Sequence 21902, A
	36	35.2	70.4	304	19	US-10-690-658-15	Sequence 15, Appl
	37	35.2	70.4	310	14	US-10-165-857-1	Sequence 1, Appl
	38	35.2	70.4	310	16	US-10-244-142A-5	Sequence 5, Appl
	39	35.2	70.4	310	17	US-10-165-764B-1	Sequence 1, Appl
	40	35.2	70.4	310	18	US-10-165-856A-1	Sequence 1, Appl1
	41	35.2	70.4	335	9	US-09-924-035A-179	Sequence 179, App
	42	35.2	70.4	337	9	US-09-969-617-4	Sequence 4, Appl
C	43	35.2	70.4	344	9	US-09-924-035A-864	Sequence 864, App
	44	35.2	70.4	351	16	US-10-244-142A-4	Sequence 4, Appl
	45	35.2	70.4	352	14	US-10-165-857-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-896-888A-10
; Sequence 10, Application US/09896888A
; Patent No. US20020116723A1
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/896,888A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/048,911
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/049,946
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-896-888A-10

Query Match 100.0%; Score 50; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 ACTTAAGCTTATAGCATGACTGCCCGCTTTCCAGTCGGGAACCTGTGCG 50
Db 1 ACTTAAGCTTATAGCATGACTGCCCGCTTTCCAGTCGGGAACCTGTGCG 50

RESULT 2


```
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 80752
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_80346C.1
US-10-437-963-80752

Query Match          71.2%; Score 35.6; DB 18; Length 415;
Best Local Similarity 82.0%; Pred. No. 4.4e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ACTTAAGCTTATAGCGATGACTGCCGCTTTCCAGTCGGGAAACCTGTGCG 50
DB 148 AATCATGTCATAGCTGTTCTGCCCGCTTTCAGTCGGGAAACCTGTGCG 99

.RESULT 6
US-09-814-353-13142
; Sequence 13142, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13142
; LENGTH: 586
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-13142

Query Match          71.2%; Score 35.6; DB 10; Length 586;
Best Local Similarity 82.0%; Pred. No. 4.6e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ACTTAAGCTTATAGCGATGACTGCCGCTTTCCAGTCGGGAAACCTGTGCG 50
DB 516 ATTTAATTCGTTGCGTCACTGCCGCTTTCAGTCGGGAAACCTGTGCG 565

.RESULT 7
US-10-027-632-2086/c
; Sequence 2086, Application US/10027632
```

```
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2086
; LENGTH: 865
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(865)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-2086

Query Match          71.2%; Score 35.6; DB 13; Length 865;
Best Local Similarity 82.0%; Pred. No. 5e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ACTTAAGCTTATAGCGATGACTGCCGCTTTCCAGTCGGGAAACCTGTGCG 50
DB 280 AATTAATTCGTTGCGTCACTGCCGCTTTCAGTCGGGAAACCTGTGCG 231

.RESULT 8
US-10-027-632-2087/c
; Sequence 2087, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2087
; LENGTH: 865
; TYPE: DNA
```

```
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(865)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-2087

Query Match      71.2%; Score 35.6; DB 13; Length 865;
Best Local Similarity 82.0%; Pred. No. 5e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 50
Db 280 AATTAAATTCGGTGGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 231

RESULT 9
US-10-027-632-2086/c
; Sequence 2086, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2086
; LENGTH: 865
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(865)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-2086

Query Match      71.2%; Score 35.6; DB 17; Length 865;
Best Local Similarity 82.0%; Pred. No. 5e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 50
Db 280 AATTAAATTCGGTGGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 231

RESULT 10
US-10-027-632-2087/c
; Sequence 2087, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
```

```
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2087
; LENGTH: 865
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(865)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-2087

Query Match      71.2%; Score 35.6; DB 17; Length 865;
Best Local Similarity 82.0%; Pred. No. 5e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 50
Db 280 AATTAAATTCGGTGGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 231

RESULT 11
US-10-437-963-88806/c
; Sequence 88806, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 88806
; LENGTH: 2307
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_87621C.1
US-10-437-963-88806

Query Match      71.2%; Score 35.6; DB 18; Length 2307;
Best Local Similarity 82.0%; Pred. No. 6e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 50
Db 120 AATTAAATTCGGTGGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 71

RESULT 12
US-10-437-963-39259/c
```

; Sequence 39259, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazov, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 39259
; LENGTH: 2721
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_42817C.1
US-10-437-963-39259

Query Match 71.2%; Score 35.6; DB 18; Length 2721;
Best Local Similarity 82.0%; Pred. No. 6.2e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTTCAGTCGGGAACCTGTGC 50
DB 120 AATCATGGTCATAGCTGTTTCTGCCCGCTTTCAGTCGGGAACCTGTGC 71

RESULT 13
US-10-357-930-36271
; Sequence 36271, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endegze, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36271
; LENGTH: 173
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-36271

Query Match 70.4%; Score 35.2; DB 18; Length 173;
Best Local Similarity 83.3%; Pred. No. 5.5e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTTCAGTCGGGAACCTGTGC 50
DB 120 AATCATGGTCATAGCTGTTTCTGCCCGCTTTCAGTCGGGAACCTGTGC 71

QY 3 TTAAGCTTATAGCGATGACTGCCCGCTTTCAGTCGGGAACCTGTGC 50
DB 63 TTAATTGCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAACCTGTGC 110

RESULT 14
US-10-357-930-45266
; Sequence 45266, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endegze, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45266
; LENGTH: 179
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-45266

Query Match 70.4%; Score 35.2; DB 18; Length 179;
Best Local Similarity 83.3%; Pred. No. 5.5e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGCGATGACTGCCCGCTTTCAGTCGGGAACCTGTGC 50
DB 68 TTAATTGCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAACCTGTGC 115

RESULT 15
US-09-764-846-347
; Sequence 347, Application US/09764846
; Patent No. US20020102638A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT212
; CURRENT APPLICATION NUMBER: US/09/764,846
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 347
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (159)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 04:33:07 ; Search time 413.906 Seconds
(without alignments)
4931.434 Million cell updates/sec

Title: US-09-896-888A-10
Perfect score: 50
Sequence: 1 acttaagcttatgcgtga.....ttccagtcggaaacctgtcg 50

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/PTUS1_COMB.seq:*

2: /cgn2_6/ptodata/1/pna/PTUS2_COMB.seq:*

3: /cgn2_6/ptodata/1/pna/PTUS3_COMB.seq:*

4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*

5: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*

6: /cgn2_6/ptodata/1/pna/US08_COMB.seq:*

7: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*

8: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*

9: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*

10: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*

11: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*

12: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*

13: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*

14: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*

15: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*

16: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*

17: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*

18: /cgn2_6/ptodata/1/pna/US092A_COMB.seq:*

19: /cgn2_6/ptodata/1/pna/US092B_COMB.seq:*

20: /cgn2_6/ptodata/1/pna/US093A_COMB.seq:*

21: /cgn2_6/ptodata/1/pna/US093B_COMB.seq:*

22: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*

23: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*

24: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*

25: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*

26: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*

27: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*

28: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*

29: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*

30: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*

31: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*

32: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*

33: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*

34: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*

35: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*

36: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*

37: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*

38: /cgn2_6/ptodata/1/pna/US098D_COMB.seq:*

39: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*

40: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*

41: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*

42: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*

43: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:*

44: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:*

45: /cgn2_6/ptodata/1/pna/US099G_COMB.seq:*

46: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*

47: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*

48: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*

49: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*

50: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*

51: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*

52: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*

53: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*

54: /cgn2_6/ptodata/1/pna/US104A_COMB.seq:*

55: /cgn2_6/ptodata/1/pna/US104B_COMB.seq:*

56: /cgn2_6/ptodata/1/pna/US105A_COMB.seq:*

57: /cgn2_6/ptodata/1/pna/US105B_COMB.seq:*

58: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*

59: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*

60: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*

61: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*

62: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*

63: /cgn2_6/ptodata/1/pna/US108A_COMB.seq:*

64: /cgn2_6/ptodata/1/pna/US108B_COMB.seq:*

65: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*

66: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*

67: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*

68: /cgn2_6/ptodata/1/pna/US110_COMB.seq:*

69: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*

70: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*

71: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*

72: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*

73: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*

74: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*

75: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*

76: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*

77: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*

78: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*

79: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*

80: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*

81: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*

82: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*

83: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*

84: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*

85: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*

86: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*

87: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*

88: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*

89: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*

90: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*

91: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*

92: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq:*

93: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq:*

94: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*

95: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*

96: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*

97: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*

98: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*

99: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*

100: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*

101: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*

102: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*

103: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*

104: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*

105: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*

106: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*

107: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*

108: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*

109: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*

110: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*

111: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*

112: /cgn2_6/ptodata/1/pna/US6042_COMB.seq:*

113: /cgn2_6/ptodata/1/pna/US6043_COMB.seq:*

114: /cgn2_6/ptodata/1/pna/US6044_COMB.seq:*

115: /cgn2_6/ptodata/1/pna/US6045_COMB.seq:*

116: /cgn2_6/ptodata/1/pna/US6046_COMB.seq:*

117: /cgn2_6/prodata/1/pna/US6047_COMB.seq.*
 118: /cgn2_6/prodata/1/pna/US6048_COMB.seq.*
 119: /cgn2_6/prodata/1/pna/US6049_COMB.seq.*
 120: /cgn2_6/prodata/1/pna/US6050_COMB.seq.*
 121: /cgn2_6/prodata/1/pna/US6051_COMB.seq.*
 122: /cgn2_6/prodata/1/pna/US6052_COMB.seq.*
 123: /cgn2_6/prodata/1/pna/US6053_COMB.seq.*
 124: /cgn2_6/prodata/1/pna/US6054_COMB.seq.*
 125: /cgn2_6/prodata/1/pna/US6055_COMB.seq.*
 126: /cgn2_6/prodata/1/pna/US6056_COMB.seq.*
 127: /cgn2_6/prodata/1/pna/US6057_COMB.seq.*
 128: /cgn2_6/prodata/1/pna/US6058_COMB.seq.*
 129: /cgn2_6/prodata/1/pna/US6059_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	50	16	US-09-048-911-10
2	50	100.0	50	38	US-09-896-888-10
3	50	100.0	50	38	US-09-896-888A-10
C 4	36	72.0	16091	1	PCT-US01-08631-25782
5	36	72.0	20795	1	PCT-US01-08631-28400
C 6	35.8	71.6	745	1	PCT-US00-00903-37
C 7	35.8	71.6	745	27	US-09-618-150-37
C 8	35.8	71.6	745	45	US-09-985-153-37
C 9	35.8	71.6	745	48	US-10-100-683-981
C 10	35.8	71.6	745	68	US-11-001-793-981
11	35.8	71.6	791	49	US-10-198-846-7241
12	35.8	71.6	869	49	US-10-198-846-4171
C 13	35.6	71.2	415	54	US-10-437-963-80752
C 14	35.6	71.2	586	35	US-09-814-353-13142
C 15	35.6	71.2	586	35	US-09-814-353A-13142
C 16	35.6	71.2	865	28	US-09-634-306B-2086
C 17	35.6	71.2	865	28	US-09-634-306B-2087
C 18	35.6	71.2	865	46	US-10-027-632-2086
C 19	35.6	71.2	865	46	US-10-027-632-2087
C 20	35.6	71.2	999	1	PCT-US01-08656-1974
C 21	35.6	71.2	999	51	US-10-273-573-1974
C 22	35.6	71.2	1637	1	PCT-US01-08631-13360
C 23	35.6	71.2	1695	1	PCT-US01-08631-13349
C 24	35.6	71.2	1942	1	PCT-US01-08631-4942
C 25	35.6	71.2	2307	54	US-10-437-963-88806
C 26	35.6	71.2	2346	1	PCT-US01-08631-13351
C 27	35.6	71.2	2424	1	PCT-US01-08631-23327
C 28	35.6	71.2	2710	1	PCT-US01-08631-13365
C 29	35.6	71.2	2710	1	PCT-US02-3955A-2277
C 30	35.6	71.2	2721	54	US-10-437-963-39259
C 31	35.6	71.2	2757	1	PCT-US01-08631-5677
C 32	35.6	71.2	5909	1	PCT-US01-08631-5236
C 33	35.6	71.2	10771	1	PCT-US01-08631-21710
C 34	35.6	71.2	16122	1	PCT-US01-14827-2281
C 35	35.6	71.2	16122	26	US-09-577-408-3214
C 36	35.6	71.2	20974	1	PCT-US01-08631-28399
C 37	35.2	70.4	118	8	US-08-221-804-1
C 38	35.2	70.4	118	10	US-08-443-781-1
C 39	35.2	70.4	134	22	US-09-426-293A-52
C 40	35.2	70.4	134	22	US-09-426-293B-52
C 41	35.2	70.4	151	7	US-08-123-456-200
C 42	35.2	70.4	151	19	US-09-297-477A-200
C 43	35.2	70.4	151	45	US-09-994-404-200
C 44	35.2	70.4	173	34	US-09-785-276A-36271
C 45	35.2	70.4	173	53	US-10-357-930-36271

ALIGNMENTS

RESULT 1
 US-09-048-911-10
 ; Sequence 10, Application US/09048911
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of British Columbia
 ; TITLE OF INVENTION: Insect Expression Vectors
 ; FILE REFERENCE: 80021-44
 ; CURRENT APPLICATION NUMBER: US/09/048,911
 ; CURRENT FILING DATE: 1998-03-26
 ; EARLIER APPLICATION NUMBER: US 60/049,946
 ; EARLIER FILING DATE: 1997-03-27
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 10
 ; LENGTH: 50
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Primer
 US-09-048-911-10
 Query Match 100.0%; Score 50; DB 16; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.9e-10;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTCCAGTCGGGAACCTGTGCG 50
 DB 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTCCAGTCGGGAACCTGTGCG 50
 RESULT 2
 US-09-896-888-10
 ; Sequence 10, Application US/09896888
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of British Columbia
 ; TITLE OF INVENTION: Insect Expression Vectors
 ; FILE REFERENCE: 80021-44
 ; CURRENT APPLICATION NUMBER: US/09/896,888
 ; CURRENT FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: US 09/048,911
 ; PRIOR FILING DATE: 1998-03-26
 ; PRIOR APPLICATION NUMBER: US 60/049,946
 ; PRIOR FILING DATE: 1997-03-27
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 10
 ; LENGTH: 50
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Primer
 US-09-896-888-10
 Query Match 100.0%; Score 50; DB 38; Length 50;
 Best Local Similarity 100.0%; Pred. No. 2.9e-10;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTCCAGTCGGGAACCTGTGCG 50
 DB 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTCCAGTCGGGAACCTGTGCG 50
 RESULT 3
 US-09-896-888A-10
 ; Sequence 10, Application US/09896888A
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of British Columbia
 ; TITLE OF INVENTION: Insect Expression Vectors
 ; FILE REFERENCE: 80021-44
 ; CURRENT APPLICATION NUMBER: US/09/896,888A
 ; CURRENT FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: US/09/048,911
 ; PRIOR FILING DATE: 1998-03-26

```

/ NAME/KEY: SITE
/ LOCATION: (113)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (163)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (727)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (739)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (745)

```

; OTHER INFORMATION: n equals a,t,g, or c
PCT-US00-00903-37

Query Match 71.6%; Score 35.8; DB 1; Length 745;
Best Local Similarity 83.3%; Pred. No. 0.00049;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 3 TTAAGCTTATAGCATGACTGCCCGCTTTCCAGTCGGGAACCTGTGCG 50
Db 168 TTAATNGCGTTCGCTCACTGCCCGCTTTCCAGTCGGGAACCTGTGCG 121

RESULT 7

US-09-618-150-37/c

; Sequence 37, Application US/09618150

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 33 Human Secreted Proteins

; FILE REFERENCE: P2036P1

; CURRENT APPLICATION NUMBER: US/09/618,150

; PRIOR FILING DATE: 2000-07-17

; PRIOR APPLICATION NUMBER: PCT/US00/00903

; PRIOR FILING DATE: 2000-01-18

; PRIOR APPLICATION NUMBER: 60/116,330

; PRIOR FILING DATE: 1999-01-19

; NUMBER OF SEQ ID NOS: 147

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 37

; LENGTH: 745

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (3)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (27)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (48)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (93)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (113)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (163)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (727)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (739)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (745)

; OTHER INFORMATION: n equals a,t,g, or c

; US-09-618-150-37

Query Match 71.6%; Score 35.8; DB 27; Length 745;

Best Local Similarity 83.3%; Pred. No. 0.00049;

Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 3 TTAAGCTTATAGCATGACTGCCCGCTTTCCAGTCGGGAACCTGTGCG 50
Db 168 TTAATNGCGTTCGCTCACTGCCCGCTTTCCAGTCGGGAACCTGTGCG 121

RESULT 8

US-09-985-153-37/c

; Sequence 37, Application US/09985153

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 33 Human Secreted Proteins

; FILE REFERENCE: P2036P1

; CURRENT APPLICATION NUMBER: US/09/985,153

; CURRENT FILING DATE: 2001-11-01

; PRIOR APPLICATION NUMBER: 09/618,150

; PRIOR FILING DATE: 2000-07-17

; PRIOR APPLICATION NUMBER: PCT/US00/00903

; PRIOR FILING DATE: 2000-01-18

; PRIOR APPLICATION NUMBER: 60/116,330

; PRIOR FILING DATE: 1999-01-19

; NUMBER OF SEQ ID NOS: 147

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 37

; LENGTH: 745

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (3)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (27)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (48)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (93)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (113)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (163)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (727)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (739)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (745)

; OTHER INFORMATION: n equals a,t,g, or c

; US-09-985-153-37

Query Match 71.6%; Score 35.8; DB 45; Length 745;

Best Local Similarity 83.3%; Pred. No. 0.00049;

Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 3 TTAAGCTTATAGCATGACTGCCCGCTTTCCAGTCGGGAACCTGTGCG 50
Db 168 TTAATNGCGTTCGCTCACTGCCCGCTTTCCAGTCGGGAACCTGTGCG 121

RESULT 9

US-10-100-683-981/c

; Sequence 981, Application US/10100683

; GENERAL INFORMATION:

; APPLICANT: Rosen, et al.

; TITLE OF INVENTION: Human Secreted Proteins

; FILE REFERENCE: P8900

; CURRENT APPLICATION NUMBER: US/10/100,683

; CURRENT FILING DATE: 2002-03-19

; PRIOR APPLICATION NUMBER: US 60/040,162

; PRIOR FILING DATE: 1997-03-07

; PRIOR APPLICATION NUMBER: US 60/043,576

; PRIOR FILING DATE: 1997-04-11

; PRIOR APPLICATION NUMBER: US 60/047,601

; PRIOR FILING DATE: 1997-05-23

; PRIOR APPLICATION NUMBER: US 60/056,845

```

; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,599
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,892
; PRIOR FILING DATE: 1997-08-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13468
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 981
; LENGTH: 745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3)..(13)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (27)..(27)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (48)..(48)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (93)..(93)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (113)..(113)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163)..(163)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (727)..(727)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (739)..(739)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (745)..(745)
; OTHER INFORMATION: n equals a,t,g, or c
;
; US-10-100-683-981
;
; Query Match 71.6%; Score 35.8; DB 48; Length 745;
; Best Local Similarity 83.3%; Pred.No. 0.00049;
; Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

Qy

3 TTAAGCTTATAGCGATGACTGCCCGCTTTCCAGTCGGGAACCTGTGC 50
||||| |||| | ||||| ||||| ||||| ||||| ||||| ||||| |||||

Dβ

168 TTAAATNGC GTTGCGTCACTGCCCGCTTTCCAGTCGGGAACCTGTGC 121

```

RESULT 10
US-11-001-793-981/c
; Sequence 981: Application US/11001793
; GENERAL INFORMATION:
; APPLICANT: Rosen, et al.
; TITLE OF INVENTION: Human secreted proteins

```

```

; FILE REFERENCE: PS900
; CURRENT APPLICATION NUMBER: US/11/001,793
; CURRENT FILING DATE: 2004-12-02
; PRIOR APPLICATION NUMBER: US/10/100,683
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: US 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,599
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,632
; PRIOR FILING DATE: 1997-05-23
; Remaining Prior Application data removed
; NUMBER OF SEQ ID NOS: 13468
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 981
; LENGTH: 745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3)..(3)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (27)..(27)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (48)..(48)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (93)..(93)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (113)..(113)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (163)..(163)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (727)..(727)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (739)..(739)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (745)..(745)
; OTHER INFORMATION: n equals a,t,g, or c
; US-11-001-793-981

```

Query Match 71.6%; Score 35.8; DB 68; Length 745;
Best Local Similarity 83.3%; Pred. No. 0.00049;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 168 TTAATNGCGTTCGGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCG 121
|||||

RESULT 11

US-10-198-846-7241
; Sequence 7241, Application US/10198846
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7241
; LENGTH: 791
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: 282, 300, 554, 592, 650, 707, 720, 728, 755, 767, 775, 791
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-7241

Query Match 71.6%; Score 35.8; DB 49; Length 791;
Best Local Similarity 83.3%; Pred. No. 0.0005;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 3 TTAACCTTATAGCGATGACTGCCCGCTTTCCAGTCGGGAAACCTGTCG 50
|||||

Db 544 TTAATTGCGTNGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCG 591
|||||

RESULT 12

US-10-198-846-4171
; Sequence 4171, Application US/10198846
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4171
; LENGTH: 869
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: 2, 3, 4, 6, 9, 10, 11, 531, 585, 667, 691, 696, 732, 742,
; LOCATION: 763, 764, 780, 816, 841, 856
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-4171

Query Match 71.6%; Score 35.8; DB 49; Length 869;
Best Local Similarity 83.3%; Pred. No. 0.00051;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 3 TTAACCTTATAGCGATGACTGCCCGCTTTCCAGTCGGGAAACCTGTCG 50
|||||

Db 569 TTAATTGCGTTCGGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCG 616
|||||

RESULT 13

US-10-437-963-80752/c
; Sequence 80752, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 80752
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_80346C.1
US-10-437-963-80752

Query Match 71.2%; Score 35.6; DB 54; Length 415;
Best Local Similarity 82.0%; Pred. No. 0.00054;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTTCCAGTCGGGAAACCTGTCG 50
|||||

Db 148 AATCATGTCATAGCTGTTCTGCGCGCTTTCCAGTCGGGAAACCTGTCG 99
|||||

RESULT 14

US-09-814-353-13142
; Sequence 13142, Application US/09814353
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13142
; LENGTH: 586
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-13142

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 05:51:35 ; Search time 123.262 Seconds
(without alignments)
834.449 Million cell updates/sec

Title: US-09-896-888A-10
Perfect score: 50
Sequence: 1 acttaagcttatagcagatga.....tccagtcgggaacctgtcg 50

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9398789 seqs, 102855566 residues

Total number of hits satisfying chosen parameters: 18797578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_NA_New.*

- 1: /cgn2_6/prodata/1/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/prodata/1/pna/PCT_NEW_COMB.seq2.*
- 3: /cgn2_6/prodata/1/pna/US05_NEW_COMB.seq.*
- 4: /cgn2_6/prodata/1/pna/US07_NEW_COMB.seq.*
- 5: /cgn2_6/prodata/1/pna/US08_NEW_COMB.seq.*
- 6: /cgn2_6/prodata/1/pna/US09_NEW_COMB.seq.*
- 7: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq.*
- 8: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq2.*
- 9: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq3.*
- 10: /cgn2_6/prodata/1/pna/US11_NEW_COMB.seq.*
- 11: /cgn2_6/prodata/1/pna/US11_NEW_COMB.seq2.*
- 12: /cgn2_6/prodata/1/pna/US11_NEW_COMB.seq3.*
- 13: /cgn2_6/prodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36.8	73.6	451	9	US-10-703-032-102508
2	36	72.0	16091	8	US-10-450-763-25782
3	36	72.0	20795	8	US-10-450-763-28400
4	35.6	71.2	1637	8	US-10-450-763-13360
5	35.6	71.2	1695	8	US-10-450-763-13349
6	35.6	71.2	1942	8	US-10-450-763-4942
7	35.6	71.2	2346	8	US-10-450-763-13351
8	35.6	71.2	2424	8	US-10-450-763-23327
9	35.6	71.2	2710	8	US-10-450-763-13365
10	35.6	71.2	2757	8	US-10-450-763-5677
11	35.6	71.2	5909	8	US-10-450-763-5236
12	35.6	71.2	10771	8	US-10-450-763-21710
13	35.6	71.2	20974	8	US-10-450-763-28399
14	35.2	70.4	271	11	US-11-035-627-7
15	35.2	70.4	282	8	US-10-450-763-21773
16	35.2	70.4	364	9	US-10-472-963-1746
17	35.2	70.4	364	9	US-10-472-963-1748
18	35.2	70.4	389	8	US-10-450-763-6994
19	35.2	70.4	389	8	US-10-450-763-10137
20	35.2	70.4	395	9	US-10-841-603B-20

21	35.2	70.4	421	10	US-11-090-997-2544	Sequence 2544, Ap
22	35.2	70.4	456	10	US-11-090-997-2578	Sequence 2578, Ap
c 23	35.2	70.4	473	8	US-10-450-763-18863	Sequence 18863, A
24	35.2	70.4	481	2	PCT-US05-06626-428	Sequence 428, App
25	35.2	70.4	508	2	PCT-US05-06626-55	Sequence 55, Appl
26	35.2	70.4	508	9	US-10-703-032-102521	Sequence 102521, A
c 27	35.2	70.4	519	8	US-10-450-763-7589	Sequence 7589, Ap
c 28	35.2	70.4	519	8	US-10-450-763-13803	Sequence 13803, A
29	35.2	70.4	526	2	PCT-US05-06626-54	Sequence 54, Appl
30	35.2	70.4	529	2	PCT-US05-06626-7	Sequence 7, Appl
31	35.2	70.4	540	11	US-11-022-102-2	Sequence 2, Appl
32	35.2	70.4	536	13	US-60-655-875-85282	Sequence 85282, A
c 33	35.2	70.4	634	8	US-10-450-763-18949	Sequence 18949, A
c 34	35.2	70.4	665	13	US-60-655-875-88301	Sequence 88301, A
c 35	35.2	70.4	686	9	US-10-472-963-1156	Sequence 1156, Ap
c 36	35.2	70.4	686	9	US-10-472-963-1747	Sequence 1747, Ap
37	35.2	70.4	690	10	US-11-090-997-2551	Sequence 2551, Ap
38	35.2	70.4	705	2	PCT-US05-06626-524	Sequence 524, App
39	35.2	70.4	717	9	US-10-703-032-74322	Sequence 74322, A
40	35.2	70.4	730	8	US-10-450-763-9465	Sequence 9465, Ap
41	35.2	70.4	782	13	US-60-655-875-88287	Sequence 88287, A
c 42	35.2	70.4	793	8	US-10-450-763-10316	Sequence 10316, A
43	35.2	70.4	854	9	US-10-703-032-104007	Sequence 104007, A
c 44	35.2	70.4	861	8	US-10-450-763-28930	Sequence 28930, A
45	35.2	70.4	868	2	PCT-US05-06626-551	Sequence 551, App

ALIGNMENTS

RESULT 1

US-10-703-032-102508

; Sequence 102508, Application US/10703032

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; APPLICANT: Andersen, Scott E.

; APPLICANT: Byrum, Joseph R.

; APPLICANT: Conner, Timothy W.

; APPLICANT: Cao, Yongwei

; APPLICANT: Masucci, James D.

; APPLICANT: Zhou, Yihua

; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53374)B

; CURRENT APPLICATION NUMBER: US/10/703,032

; CURRENT FILING DATE: 2003-11-06

; PRIOR APPLICATION NUMBER: 10/020,338

; PRIOR FILING DATE: 2001-12-12

; NUMBER OF SEQ ID NOS: 211164

; SEQ ID NO 102508

; LENGTH: 451

; TYPE: DNA

; ORGANISM: Triticum aestivum

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_TA_102508

US-10-703-032-102508

Query Match 73.6%; Score 36.8; DB 9; Length 451;

Best Local Similarity 85.4%; Pred. No. 8e-06; 7; Indels 0; Gaps 0;

Matches 41; Conservative 0; Mismatches 0;

QY 3 TTAAGCTTATAGCGATGCTCCGCTTTCCAGTCGGGAACCTGTGC 50

|||||

Db 94 TTAAATTCGTTGGCTCACTGCCGCTTTCCAGTCGGGAACCTGTGC 141

RESULT 2

US-10-450-763-25782/c

; Sequence 25782, Application US/10450763

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

; FILE REFERENCE: 790CIP3/US

RESULT 4
US-10-450-763-13360/c

Query Match	71.2%;	Score 35.6;	DB 8;
Best Local Similarity	82.0%;	Pred. No. 3.4e-05;	Length 1695;

RESULT 4
US-10-450-763-13360/c

Best Local Similarity 82.0%; Pred. No. 3.6e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ACTTAAGCTTATAGCGNATGACTGCCGCTTTTCAGTCGGGAACCTGTG 50
|||||
Db 1672 AATTAAATGGCTTGGCTGCTACTGCCGCTTTTCAGTCGGGAACCTGTG 1721

```

RESULT 8
US-10-450-763-23327/c
; Sequence 23327, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450.763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 23327
; LENGTH: 2424
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1801)..(2265)
; OTHER INFORMATION: 100% homologous to Homo sapiens A human prolif
; OTHER INFORMATION: apoptosis related protein,accession number Y84
; OTHER INFORMATION: Score=838.
US-10-450-763-23327

```

```

Qy      1  ACTTAAAGCTTATACGGATGACTGCCCGCTTTCCAGTCGGGAAACCTGTGC  50
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     1183  AATTAAATTGGCTTCGCGTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGC  1134

```

```

RESULT 9
US-10-450-763-13365/c
; Sequence 13365, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 13365
; LENGTH: 2710
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1)..(228)
; OTHER INFORMATION: 94% homologous to Homo sapiens Alzheimer's dis
; OTHER INFORMATION: encoded by DNA from plasmid pGCS22322.accession
; OTHER INFORMATION: Waterman Score=392.

```

US-10-450-763-13365

Query Match 71.2%; Score 35.6; DB 8; Length 2710;
Best Local Similarity 82.0%; Pred. No. 3.7e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 50
Db 1080 AATTAATGGCTTGGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 1031

RESULT 10

US-10-450-763-5677/c
; Sequence 5677, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 5677
; LENGTH: 2757
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1971)..(2030)
; OTHER INFORMATION: 50% homologous to Cloning vector pZC320 SopA, accession number
; OTHER INFORMATION: U26464,Smith-Waterman Score=40.

US-10-450-763-5677

Query Match 71.2%; Score 35.6; DB 8; Length 2757;
Best Local Similarity 82.0%; Pred. No. 3.7e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 50
Db 675 AATTAATGGCTTGGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 626

RESULT 11

US-10-450-763-5236/c
; Sequence 5236, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 5236
; LENGTH: 5909
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (5001)..(5864)
; OTHER INFORMATION: 100% homologous to Cloning vector pZC320 SopA, accession

; OTHER INFORMATION: number U26464,Smith-Waterman Score=1509.
US-10-450-763-5236

Query Match 71.2%; Score 35.6; DB 8; Length 5909;
Best Local Similarity 82.0%; Pred. No. 4.3e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 50
Db 737 AATTAATGGCTTGGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 688

RESULT 12

US-10-450-763-21710/c
; Sequence 21710, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 21710
; LENGTH: 10771
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1)..(840)
; OTHER INFORMATION: 94% homologous to Cloning vector pZC320 SopA, accession number
; OTHER INFORMATION: U26464,Smith-Waterman Score=1403.

US-10-450-763-21710

Query Match 71.2%; Score 35.6; DB 8; Length 10771;
Best Local Similarity 82.0%; Pred. No. 4.8e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ACTTAAGCTTATAGCGATGACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 50
Db 10060 AATTAATGGCTTGGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGCG 10011

RESULT 13

US-10-450-763-28399/c
; Sequence 28399, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 28399
; LENGTH: 20974
; TYPE: DNA

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (7)..(387)
; OTHER INFORMATION: 100% homologous to Bacillus amyloliquefaciens precursor (AA -
; OTHER INFORMATION: 29 to 443), accession number X52988, Smith-Waterman Score=664.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(20974)
; OTHER INFORMATION: n = a,t,c or g
US-10-450-763-28399
```

```
Query Match      71.2%; Score 35.6; DB 8; Length 20974;
Best Local Similarity 82.0%; Pred. No. 5.5e-05;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY      1 ACTTAAGCTTATAGCGATGACTGCCCGCTTTCAGTCGGGAACCTGTGCG 50
Db      3214 AATTAATGCGTGGCTCACTGCCCGCTTTCAGTCGGGAACCTGTGCG 3165
```

```
RESULT 14
US-11-035-627-7/c
; GENERAL INFORMATION:
; SEQUENCE 7, Application US/11035627
; APPLICANT: University of North Carolina at Chapel Hill
; APPLICANT: Wilson, W. David
; APPLICANT: Boykin, David W
; APPLICANT: Tidwell, Richard R
; TITLE OF INVENTION: NOVEL COMPOUNDS THAT EXHIBIT SPECIFIC MOLECULAR RECOGNITION OF
; TITLE OF INVENTION: MIXED NUCLEIC ACID SEQUENCES AND BIND IN THE DNA MINOR GROOVE AS
; FILE REFERENCE: 421/60/16/2/2/2
; CURRENT APPLICATION NUMBER: US/11/035,627
; PRIOR FILING DATE: 2005-01-15
; PRIOR APPLICATION NUMBER: US 10/653,677
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 09/745,004
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: US 60/172,863
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 271
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: EcrR1-PvuII fragment from plasmid pBS+, antisense strand
US-11-035-627-7
```

```
Query Match      70.4%; Score 35.2; DB 11; Length 271;
Best Local Similarity 83.3%; Pred. No. 3.5e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY      3 TTAAGCTTATAGCGATGACTGCCCGCTTTCAGTCGGGAACCTGTGCG 50
Db      57 TTAATTCGTTGGCTCACTGCCCGCTTTCAGTCGGGAACCTGTGCG 10
```

```
RESULT 15
US-10-450-763-21773/c
; SEQUENCE 21773, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
```

```
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 21773
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (243)..(214)
; OTHER INFORMATION: 100% homologous to Homo sapiens amino acid sequence of a
; OTHER INFORMATION: human secreted peptide, accession number Y10930, Smith-Waterman Sco
```

```
US-10-450-763-21773
```

```
Query Match      70.4%; Score 35.2; DB 8; Length 282;
Best Local Similarity 83.3%; Pred. No. 3.5e-05;
Matches 40; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY      3 TTAAGCTTATAGCGATGACTGCCCGCTTTCAGTCGGGAACCTGTGCG 50
Db      114 TTAATTCGTTGGCTCACTGCCCGCTTTCAGTCGGGAACCTGTGCG 67
```

```
Search completed: May 10, 2005, 09:51:26
Job time : 123.262 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 03:45:00 ; Search time 350.316 Seconds
(without alignments)
5432.847 Million cell updates/sec

Title: US-09-896-888A-10
Perfect score: 50
Sequence: 1 acttaagcttatagcatga.....tccagtcgggaacctgtcg 50
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38.4	76.8	412	9	CNS07H0H
2	36.8	73.6	125	8	AQ080570 CIT-HSP-2
3	36.8	73.6	174	8	AQ076128 CIT-HSP-2
4	36.8	73.6	177	8	AQ076155 CIT-HSP-2
5	36.8	73.6	311	8	AQ076114 CIT-HSP-2
6	36.8	73.6	372	7	CO990192 UMC-pd3ov
7	36.8	73.6	373	7	CO947892 UMC-p8mm2
8	36.8	73.6	425	7	CO954011 UMC-pnata
9	36.8	73.6	451	2	BE403103 GEX002.C1
10	36.8	73.6	455	9	CNS07H8S
11	36.8	73.6	591	7	CO048749 tk66a07.b
12	36.8	73.6	641	9	AG068743 Pan trogl
13	36.8	73.6	655	9	AG055437 Pan trogl
14	36.8	73.6	681	9	AG125607 Pan trogl
15	36.8	73.6	682	9	AG118925 Pan trogl
16	36.8	73.6	723	9	AG068231 Pan trogl
17	36.8	73.6	745	9	AG030401 Pan trogl
18	36.8	73.6	755	9	AG102051 Pan trogl
19	36.8	73.6	784	9	AG383436 Mus muscu
20	36.8	73.6	885	2	BE229292 98BS0512
21	36.4	72.8	350	1	AU233787 AU233787
22	36.4	72.8	357	1	AU278188 AU278188
23	36.2	72.4	674	8	BH243517 AU1R556TF
24	35.8	71.6	232	1	AU069015 AU069015

25	35.8	71.6	331	4	BJ691809
26	35.8	71.6	423	4	BJ668962
27	35.8	71.6	459	1	AU069599
28	35.8	71.6	477	5	BP874955
29	35.8	71.6	508	4	BJ671717
30	35.8	71.6	541	5	BQ079999
31	35.8	71.6	571	4	BM419897
32	35.8	71.6	586	6	CB865576
33	35.8	71.6	675	9	AG060720 Pan trogl
34	35.8	71.6	699	9	AG043113 Pan trogl
35	35.8	71.6	746	9	AG134861 Pan trogl
36	35.8	71.6	789	9	ATH517143
37	35.8	71.6	817	7	CF569104
38	35.6	71.2	368	4	B1937595
39	35.6	71.2	410	8	B30742
40	35.6	71.2	496	2	BF703023
41	35.6	71.2	692	9	AG109195
42	35.6	71.2	719	9	AG045821 Pan trogl
43	35.4	70.8	939	9	CL096240
44	35.2	70.4	117	8	AQ041632 CIT-HSP-2
45	35.2	70.4	120	9	CC961712 BOLDY76TF

ALIGNMENTS

RESULT 1
CNS07H0H 412 bp DNA linear GSS 02-OCT-2001
LOCUS Anopheles gambiae GSS T7 end of clone 23p13 of library NotreDame1
DEFINITION from strain PEST of Anopheles gambiae (African malaria mosquito), genomic survey sequence.
ACCESSION AL610451.1 GI:15916636
VERSION
KEYWORDS GSS.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
REFERENCE 1 (bases 1 to 412)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2001) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
REFERENCE 2 (bases 1 to 412)
AUTHORS Roth,C.W., Brey,P.T., Ke,Z. and Collins,F.H.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2001) BBMI, Institut Pasteur, 25, rue du Dr. Roux, Paris 75015, France
COMMENT This clone is from an A. gambiae BAC library provided by F.H. Collins and sequenced by Genoscope in collaboration with the Laboratory of Biochem. and Biol. Molec. of Insects, Institut Pasteur.

FEATURES

Location/Qualifiers
source 1..412
/organism="Anopheles gambiae"
/mol_type="genomic DNA"
/strain="PEST"
/db_xref="taxon:7165"
/clone="23p13"
/clone_lib="NotreDame1"
/note="end : T7"

ORIGIN

Query Match 76.8%; Score 38.4; DB 9; Length 412;
Best Local Similarity 87.5%; Pred. No. 0.00023;
Matches 42; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3 TTAAGCTTATAGCATGATCGCCGCTTTCAGTCGGGAAACCTGTGC 50
|||||
DB 152 TTAATTTTTCGCTCACTGCCGCTTTCAGTCGGGAAACCTGTGC 199
|||||

```

RESULT 2
AQ080570
LOCUS
DEFINITION
  CIT-HSP-2358M2.TF CIT-HSP Homo sapiens genomic clone 2358M2,
  genomic survey sequence.
ACCESSION
  AQ080570
VERSION
  AQ080570.1 GI:3441754
SOURCE
  GSS.
ORGANISM
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 125)
  Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
  Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
  Venter,J.C.
  Use of a random human BAC End Sequence Database for Sequence-Ready
  Map Building
  Unpublished (1998)
  Other_GSSs: CIT-HSP-2358M2.TR
  Contact: Mark Adams
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: mdadams@tigr.org
  Clones are available from Research Genetics (info@resgen.com). BAC
  end search page:
  http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
  Seq primer: M13-21
  Class: BAC ends.
  Location/Qualifiers
    1..125
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /clone="2358M2"
    /sex="Male"
    /cell_type="Sperm"
    /clone_lib="CIT-HSP"
    /note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
    HindII"

FEATURES
  source
    1..125
    Query Match 73.6%; Score 36.8; DB 8; Length 125;
    Best Local Similarity 85.4%; Pred. No. 0.00082;
    Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TTAAGCTTATACGATGACTGCCGCTTCCAGTCGGGAACCTGTGCG 50
    |||||
    25 TTAATTGCGTAGCGCTCACTGCCGCTTCCAGTCGGGAACCTGTGCG 72
    |||||

ORIGIN
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 174)
  Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
  Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
  Venter,J.C.
  Use of a random human BAC End Sequence Database for Sequence-Ready
  Map Building
  Unpublished (1998)
  Contact: Mark Adams
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: mdadams@tigr.org
  Clones are available from Research Genetics (info@resgen.com). BAC
  end search page:
  http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
  Seq primer: M13-21
  Class: BAC ends.
  Location/Qualifiers
    1..174
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /clone="2364K4"
    /sex="Male"
    /cell_type="Sperm"
    /clone_lib="CIT-HSP"
    /note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
    HindII"

FEATURES
  source
    1..174
    Query Match 73.6%; Score 36.8; DB 8; Length 174;
    Best Local Similarity 85.4%; Pred. No. 0.00086;
    Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TTAAGCTTATACGATGACTGCCGCTTCCAGTCGGGAACCTGTGCG 50
    |||||
    25 TTAATTGCGTAGCGCTCACTGCCGCTTCCAGTCGGGAACCTGTGCG 123
    |||||

ORIGIN
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 177)
  Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
  Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
  Venter,J.C.
  Use of a random human BAC End Sequence Database for Sequence-Ready
  Map Building
  Unpublished (1998)
  Contact: Mark Adams
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: mdadams@tigr.org
  Clones are available from Research Genetics (info@resgen.com). BAC
  end search page:
  http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
  Seq primer: M13-21
  Class: BAC ends.
  Location/Qualifiers
    1..177
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /clone="2364M4"

FEATURES
  source
    1..177
    Query Match 73.6%; Score 36.8; DB 8; Length 177;
    Best Local Similarity 85.4%; Pred. No. 0.00086;
    Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TTAAGCTTATACGATGACTGCCGCTTCCAGTCGGGAACCTGTGCG 50
    |||||
    25 TTAATTGCGTAGCGCTCACTGCCGCTTCCAGTCGGGAACCTGTGCG 123
    |||||

ORIGIN
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 177)
  Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
  Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
  Venter,J.C.
  Use of a random human BAC End Sequence Database for Sequence-Ready
  Map Building
  Unpublished (1998)
  Contact: Mark Adams
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: mdadams@tigr.org
  Clones are available from Research Genetics (info@resgen.com). BAC
  end search page:
  http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
  Seq primer: M13-21
  Class: BAC ends.
  Location/Qualifiers
    1..177
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /clone="2364M4"

```

```

JOURNAL
COMMENT
  Map Building
  Unpublished (1998)
  Contact: Mark Adams
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: mdadams@tigr.org
  Clones are available from Research Genetics (info@resgen.com). BAC
  end search page:
  http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
  Seq primer: M13-21
  Class: BAC ends.
  Location/Qualifiers
    1..174
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /clone="2364K4"
    /sex="Male"
    /cell_type="Sperm"
    /clone_lib="CIT-HSP"
    /note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
    HindII"

FEATURES
  source
    1..174
    Query Match 73.6%; Score 36.8; DB 8; Length 174;
    Best Local Similarity 85.4%; Pred. No. 0.00086;
    Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TTAAGCTTATACGATGACTGCCGCTTCCAGTCGGGAACCTGTGCG 50
    |||||
    25 TTAATTGCGTAGCGCTCACTGCCGCTTCCAGTCGGGAACCTGTGCG 123
    |||||

ORIGIN
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 177)
  Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
  Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
  Venter,J.C.
  Use of a random human BAC End Sequence Database for Sequence-Ready
  Map Building
  Unpublished (1998)
  Contact: Mark Adams
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: mdadams@tigr.org
  Clones are available from Research Genetics (info@resgen.com). BAC
  end search page:
  http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
  Seq primer: M13-21
  Class: BAC ends.
  Location/Qualifiers
    1..177
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /clone="2364M4"

FEATURES
  source
    1..177
    Query Match 73.6%; Score 36.8; DB 8; Length 177;
    Best Local Similarity 85.4%; Pred. No. 0.00086;
    Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TTAAGCTTATACGATGACTGCCGCTTCCAGTCGGGAACCTGTGCG 50
    |||||
    25 TTAATTGCGTAGCGCTCACTGCCGCTTCCAGTCGGGAACCTGTGCG 123
    |||||

ORIGIN
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 177)
  Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
  Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
  Venter,J.C.
  Use of a random human BAC End Sequence Database for Sequence-Ready
  Map Building
  Unpublished (1998)
  Contact: Mark Adams
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: mdadams@tigr.org
  Clones are available from Research Genetics (info@resgen.com). BAC
  end search page:
  http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
  Seq primer: M13-21
  Class: BAC ends.
  Location/Qualifiers
    1..177
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /clone="2364M4"

```

```

/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match      73.6%; Score 36.8; DB 8; Length 177;
Best Local Similarity 85.4%; Pred. No. 0.00086;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 3 TTAAGCTTATAGCATGACGTCGCCGCTTTCAGTCGCGGAACCTGTCG 50
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 83 TTAATAGCGTAGCGCTCACTGCCGCGCTTTCAGTCGCGGAACCTGTCG 130

RESULT 5
LOCUS AQ076114
DEFINITION CIT-HSP-2368K7.TF CIT-HSP Homo sapiens genomic clone 2368K7,
genomic survey sequence.
ACCESSION AQ076114.1 GI:3436248
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 311)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Map Building
COMMENT Unpublished (1998)
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tcdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES
source
1..311
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2368K7"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match      73.6%; Score 36.8; DB 8; Length 311;
Best Local Similarity 85.4%; Pred. No. 0.00095;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 3 TTAAGCTTATAGCATGACGTCGCCGCTTTCAGTCGCGGAACCTGTCG 50
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 232 TTAATCGCGTAGCGCTCACTGCCGCGCTTTCAGTCGCGGAACCTGTCG 279

RESULT 6
LOCUS AQ090192
DEFINITION UMC-pd3ov2-002-g02 Oviduct gilt D3 of estrous cycle pd3ov Sus


```

```

ACCESSION CO990192
VERSION CO990192.1 GI:51349466
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
REFERENCE 1 (bases 1 to 372)
AUTHORS Kim,J., Whitworth,K.M., Springer,G.K., Bivens,N.J., Ries,J.E.,
Woods,R.J., Spollen,W.G., Forrester,L.J., Mathialagan,N.,
Prather,R.S. and Green,J.A.
TITLE Large-scale Generation and Analysis of Expressed Sequence Tags from
Porcine endometrium and oviduct
JOURNAL Unpublished (2004)
COMMENT Contact: DNA Core Facility (Swine Project)
Animal Science - RS Prather
University of Missouri-Columbia
M616 Medical Sciences Bldg., Columbia, MO 65212, USA
Tel: (573)882-0428
Fax: (573)884-5552
Email: porcine@rnet.missouri.edu
POLYA-No.

```

```

FEATURES
source
1..372

```

```

/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/dev_stage="Oviduct from a gilt on day 3 of the estrous
cycle"
/clone_lib="pd3ov"
/note="Vector: pSPORT1; Funding: A grant from the Monsanto
Company to the University of Missouri. Genetic Source:
Endometrium and oviduct tissues from various stages of the
estrous cycle were collected from crossbred pigs (Sus
scrofa domestica), frozen in liquid nitrogen immediately
after collection, and stored at -80 degrees Celsius until
RNA extraction. The specific tissues collected were Day 0
and Day 3 whole oviducts and Days 3, 6, 10 and 12-14
endometrium. More information regarding the methods can be
found at:
http://genome.mnet.missouri.edu/swine/Methods.html.
Library Construction (Standard Protocol): All procedures
discussed in this section have been described in detail
elsewhere (Soares et al., 1994; Bonaldo et al., 1996;
Jiang et al., 2001). Total cellular RNA from each sample
was isolated by using STAT-60 reagent (Tel-Test,
Friendswood, TX) and poly(A)+ RNA was obtained by two
rounds of purification with the Oligotex mRNA isolation
kit (Qiagen) according to the manufacturer's instructions.
The oviduct libraries and the Day 3, 6 and 10 endometrium
libraries were constructed essentially as described by the
manufacturer's instructions provided with the SuperScript
Plasmid System (Invitrogen, cat. no. 18248-013). Briefly,
1mg of poly(A)+ RNA will be annealed at 37 degrees Celsius
with 10mg of NotI-tag-dT18 oligonucleotide
(CCTGCTCGCGCGC-tag-T18) and reverse transcribed at 37
degrees Celsius with Superscript II (Invitrogen) reverse
transcriptase (Jiang et al., 2001). The 'tag' represents a
tissue/stage-specific ten-base sequence identifier
(http://genome.uiowa.edu/pubsoft/software.html) present in
the oligonucleotide used to prime first-strand synthesis.
Second strand synthesis was performed with T4 DNA
polymerase in the presence of DNA ligase and RNase H.
After second strand synthesis, the double-stranded cDNAs
was ligated to Sali adapters (Invitrogen-Life
Technologies) and digested with NotI. The cDNAs will be
size selected by passage through cDNA size fractionation
columns (Invitrogen-Life technologies). The cDNAs derived
from each developmental stage of a particular tissue were
mixed on an equimolar basis and ligated directionally into
the NotI and Sali sites of the pSPORT1 vector
(Invitrogen). After ligation of the inserts, the plasmids

```


Jiang et al., 2001). Total cellular RNA from each sample was isolated by using STAT-60 reagent (Tel-test, Friendswood, TX) and the poly(A)⁺ RNA was obtained by two rounds of purification with the Oligotex mRNA isolation kit (Qiagen) according to the manufacturer's instructions. The libraries were constructed essentially as described by the manufacturer's instructions provided with the Superscript Plasmid System (Invitrogen, cat. no. 18248-013). Briefly, 1mg of poly(A)⁺ RNA will be annealed at 37 degrees Celsius with 10mg of NotI-tag-dw18 oligonucleotide (GCTGCTCGCGCCG-tag-r18) and reverse transcribed at 37 degrees Celsius with Superscript II (Invitrogen) reverse transcriptase (Jiang et al., 2001). The 'tag' represents a tissue/stage-specific ten-base sequence identifier

The 'tag' represents a tissue/stage-specific ten-base sequence identifier

The oligonucleotide used to prime first-strand synthesis. Second strand synthesis was performed with T4 DNA polymerase in the presence of DNA ligase and RNase H.

was ligated to SalI adapters (Invitrogen-Life Technologies) and digested with NotI. The cDNAs will be size selected by passage through cDNA size fractionation columns (Invitrogen-Life technologies). The cDNAs derived from each developmental stage of a particular tissue were mixed on an equimolar basis and ligated directionally into the NotI and SalI sites of the pSPORT1 vector (Invitrogen). After ligation of the inserts, the plasmids will be electroporated into DH10B bacteria. Preliminary Library Characterization: Randomly chosen clones from each library were analyzed by restriction digestion.

sequencing (1-4 96-well plates) to confirm library quality [e.g. the presence of short polyA+ tails, genomic DNA contamination (must be <1%), ribosomal RNA clones (must be <1%), etc.] and to provide a sequence database representing the predominant clones in each library. The

representing the predominant clones in each library. The clones were sequenced at the University of Missouri-Columbia DNA Core Facility. Bioinformatics work was performed by GK Springer's bioinformatics group (WG Spollen, JE Ries, A Guillen, AA Khambati, RV Patel, CM Topinka, SB Bhuiyan) in Computer Science and Health Management and Informatics Departments at the University of Missouri-Columbia. Clone Requests: Requests for clones should be made to the director of the University of Missouri DNA Core facility at: porcine@net.missouri.edu. Citations: 1. Bonaldo WF, Lennon G, Soares MB. *Database*

Normalization and Subtraction: Two approaches to facilitate gene discovery. Genome Res, 1996; 6:791-806.

2. Jiang H, Bivens NJ, Ries JE, Whitworth KM, Green JA, Forrester LJ, Springer GK, Didion BA, Mathialagan N, Prather RS, Lucy MC (2001) Constructing cDNA libraries with fewer clones that contain long poly(dA) tails. Biotechniques 31:38-42.

3. Soares MB, MF Bernaldo, P Jelene, L Su, L Lawton, A Efrantiadis. 1994. Construction and characterization of a normalized cDNA library. Proc Natl Acad Sci, 91:9228-9232.

TAG TISSUE=Perinatal ovary

TAG_SEQ=Not found"

73.6%; Score 36.8; DB 7; Length 425;
ty 85.4%; Pred. No. 0.001; 7; Indels 0; Gaps 0;
ervative 0; Mismatches 7; Indels 0; Gaps 0;

TTATGAGCATGACTGCCGCTTTCAGTCGGAAACCTGTGC 50
|||||
TCGTAGCGCTCACTGCCGCTTTCAGTCGGAAACCTGTGC 266

451 bp mRNA linear EST 21-JUL-2000

451 bp mRNA linear EST 21-JUL-2000

Accession	Size	Library	Library	Library
U00096	451 bp	mRNA	linear	EST 21-JUL-2000

```

DEFINITION GBX002.C11F990602#08 ITEC GBX Wheat Root Library Triticum aestivum
ACCESSION BE403103
VERSION BE403103.1 GI:9362483
SOURCE EST.
ORGANISM Triticum aestivum (bread wheat)
REFERENCE
AUTHORS Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T.,
Cloutier, S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A.,
Gustafson, P., Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J.,
Joudrier, P., Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P.,
Ogihara, Y., Pecchioni, N., Quiset, C., Schuch, W., Selvaraj, G.,
Shariflou, M., Sorrells, M., Warburton, M. and Wenzel, G.
TITLE International Triticaceae EST Cooperative (ITEC): Production of
Expressed Sequence Tags for Species of the Triticaceae
JOURNAL Unpublished (2000)
COMMENT Contact: Jacquemin JM
Centre de Recherches Agronomiques, Departement de Biotechnologie
234 chaussee de Charleroi, 5030 Gembloux BELGIUM
Tel: 32 81 61 29 35
Fax: 32 81 61 04 59
Email: jacquemin@cragx.fgov.be
International Triticaceae EST Cooperative (ITEC)
http://wheat.pw.usda.gov/genome.

FEATURES
source
location/Qualifiers
1..451
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Odeon"
/db_xref="taxon:4565"
/clone="GBX002.C11"
/dev_stage="seedling, unstressed"
/tissue_type="root"
/clone_lib="ITEC GBX Wheat Root Library"
/note="Vector: pUC18; 0.3-2.0 Kbp average insert size."

ORIGIN
Query Match 73.6%; Score 36.8; DB 2; Length 451;
Best Local Similarity 85.4%; Pred. No. 0.001;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 3 TTACCTTATAGCATGCTCCCGCTTCCAGTCGGGAACCTGTGCG 50
|||||
Db 94 TTAATTCTGCTGCTCACTGCCCGCTTCCAGTCGGGAACCTGTGCG 141

RESULT 10
CNS07H85 455 bp DNA linear GSS 02-OCT-2001
LOCUS Anopheles gambiae GSS T7 end of clone 25H23 of library Notredame1
DEFINITION from strain PEST of Anopheles gambiae (African malaria mosquito),
genomic survey sequence.
ACCESSION AL610727
VERSION AL610727.1 GI:15916912
KEYWORDS GSS.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
REFERENCE
AUTHORS Anopheles gambiae
TITLE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
1 (bases 1 to 455)
Genoscope.
DIRECT SUBMISSION
JOURNAL Submitted (01-OCT-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
REFERENCE
AUTHORS 2 (bases 1 to 455)
Roth, C.W., Brey, P.T., Ke, Z. and Collins, F.H.
TITLE Direct Submission

```

```

JOURNAL Submitted (01-OCT-2001) BEMI, Institut Pasteur, 25, rue du Dr.
Roux, Paris 75015, France
COMMENT This clone is from an A. gambiae BAC library provided by F.H.
Collins and sequenced by Genoscope in collaboration with the
Laboratory of Biochem. and Biol. Molec. of Insects, Institut
Pasteur.
FEATURES
source
location/Qualifiers
1..455
/organism="Anopheles gambiae"
/mol_type="genomic DNA"
/strain="PEST"
/db_xref="taxon:7165"
/clone="25H23"
/clone_lib="Notredame1"
/note="end : T7"

ORIGIN
Query Match 73.6%; Score 36.8; DB 9; Length 455;
Best Local Similarity 85.4%; Pred. No. 0.001;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 3 TTACCTTATAGCATGCTCCCGCTTCCAGTCGGGAACCTGTGCG 50
|||||
Db 154 TTAATTCTGCTGCTCACTGCCCGCTTCCAGTCGGGAACCTGTGCG 201

RESULT 11
CNS048749 591 bp mRNA linear EST 14-JUN-2004
LOCUS Arabidopsis RT-PCR Products (CSHL) Arabidopsis thaliana
DEFINITION tk66a07.b7 Arabidopsis RT-PCR Products (CSHL) Arabidopsis thaliana
cDNA clone tk66a07, mRNA sequence.
ACCESSION CNS048749
VERSION CNS048749.1 GI:48688371
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE
AUTHORS Katari, M.S., Balija, V. and McComb, W.R.
TITLE Arabidopsis RT-PCR Products (CSHL)
JOURNAL Unpublished (2004)
COMMENT Contact: Manpreet S. Katari
Woodbury Genome Center
Cold Spring Harbor Laboratory
500 Sunnyside Blvd., Plainview, NY 11797, USA
Tel: 516 422 4086
Fax: 516 422 4109
Email: mcombie@cshl.org
Matches At4g04920
Plate: tk66 row: a column: 07
High quality sequence stop: 591.
FEATURES
source
location/Qualifiers
1..591
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="tk66a07"
/tissue_type="Whole Plant"
/clone_lib="Arabidopsis RT-PCR Products (CSHL)"
/note="DNA was extracted from Arabidopsis thaliana whole
plant tissue, provided by members of Rob Martienssen's
lab, using TRIzol. Primers were designed in Hypothetical
genes and un-annotated regions in Arabidopsis that are
conserved in Brassica oleracea located in the short arm of
chromosome 4. PCR products were either cloned into PCR
TOPO 2.1 vector (Invitrogen) and then sequenced using -21
M13 forward and reverse universal primers or treated with
Exonuclease 1 and Shrimp Alkaline Phosphatase and
sequenced using the specific primers."

ORIGIN

```

```

Query Match      73.6%; Score 36.8; DB 7; Length 591;
Best Local Similarity 85.4%; Pred. No. 0.0011;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGGACGACGCGCGCTTCCAGTCGCGGAACCTGCG 50
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 521 TTAATGCTAGGCTCACTGCGCGCTTCCAGTCGCGGAACCTGCG 568

RESULT 12
AG068743
LOCUS      641 bp      DNA      linear      GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-059B24.F, genomic survey sequence.
ACCESSION  AG068743
VERSION     AG068743.1 GI:16620545
KEYWORDS   GSS.
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE  1
AUTHORS    Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
            Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE      BAC end sequences of Library PTB
JOURNAL    Unpublished
AUTHORS
TITLE
JOURNAL
AUTHORS
FEATURES    source
            1. .641
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            /clone="PTB-059B24.F"
            /sex="male"
            /cell_type="lymphoblast"
            /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match      73.6%; Score 36.8; DB 9; Length 641;
Best Local Similarity 85.4%; Pred. No. 0.0011;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

LIBRARY
Vector      : pKS145
R.Site 1    : SacI
R.Site 2    : SacI
Location/Qualifiers
1. .641
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-059B24.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match      73.6%; Score 36.8; DB 9; Length 655;
Best Local Similarity 85.4%; Pred. No. 0.0011;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

PRIMERS
Sequencing: -21M13
Vector      : pKS145
R.Site 1    : SacI
R.Site 2    : SacI
Location/Qualifiers
1. .655
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-041G18.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match      73.6%; Score 36.8; DB 9; Length 655;
Best Local Similarity 85.4%; Pred. No. 0.0011;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGGACGACGCGCGCTTCCAGTCGCGGAACCTGCG 50
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 TTAATGCTAGGCTCACTGCGCGCTTCCAGTCGCGGAACCTGCG 289

RESULT 14
AG125607
LOCUS      681 bp      DNA      linear      GSS 04-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-135O16.F, genomic survey sequence.
ACCESSION  AG125607
VERSION     AG125607.1 GI:16654772
KEYWORDS   GSS.
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE  1
AUTHORS    Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
            Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE      BAC end sequences of Library PTB
JOURNAL    Unpublished
AUTHORS
TITLE
JOURNAL
AUTHORS
FEATURES    source
            1. .681
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            /clone="PTB-041G18.F"
            /sex="male"
            /cell_type="lymphoblast"
            /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match      73.6%; Score 36.8; DB 9; Length 641;
Best Local Similarity 85.4%; Pred. No. 0.0011;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TTAAGCTTATAGGACGACGCGCGCTTCCAGTCGCGGAACCTGCG 50
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 217 TTAATGCTAGGCTCACTGCGCGCTTCCAGTCGCGGAACCTGCG 264

RESULT 13
AG055437
LOCUS      655 bp      DNA      linear      GSS 02-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-041G18.F, genomic survey sequence.
ACCESSION  AG055437
VERSION     AG055437.1 GI:16592880
KEYWORDS   GSS.
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE  1
AUTHORS    Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
            Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE      BAC end sequences of Library PTB
JOURNAL    Unpublished
AUTHORS
TITLE
JOURNAL
AUTHORS
FEATURES    source
            1. .655
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            /clone="PTB-041G18.F"
            /sex="male"
            /cell_type="lymphoblast"
            /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match      73.6%; Score 36.8; DB 9; Length 655;
Best Local Similarity 85.4%; Pred. No. 0.0011;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

PRIMERS
Sequencing: -21M13
Vector      : pKS145
R.Site 1    : SacI
R.Site 2    : SacI
Location/Qualifiers
1. .655
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-041G18.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match      73.6%; Score 36.8; DB 9; Length 655;
Best Local Similarity 85.4%; Pred. No. 0.0011;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 655)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
LIBRARY
Sequencing: -21M13
Vector      : pKS145
R.Site 1    : SacI
R.Site 2    : SacI
Location/Qualifiers
1. .655
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-041G18.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match      73.6%; Score 36.8; DB 9; Length 655;
Best Local Similarity 85.4%; Pred. No. 0.0011;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

PRIMERS
Sequencing: -21M13
Vector      : pKS145
R.Site 1    : SacI
R.Site 2    : SacI
Location/Qualifiers
1. .655
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-041G18.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match      73.6%; Score 36.8; DB 9; Length 655;
Best Local Similarity 85.4%; Pred. No. 0.0011;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 681)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS

```

Sequencing: -21M13

LIBRARY
Vector : PKS145
R.Site 1 : SacI
R.Site 2 : SacI

Location/Qualifiers

1..681
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-135016.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES

source

ORIGIN

Query Match 73.6%; Score 36.8; DB 9; Length 681;
Best Local Similarity 85.4%; Pred. NO. 0.0011;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;QY 3 TTAAGCTTATAGCATGCTCCGCTTCCAGTCGGGAAACCTGTGCG 50
|||||
DB 287 TTAATTGTTTGGCGCTCACTGCCGCTTCCAGTCGGGAAACCTGTGCG 334

RESULT 15

AG118925
LOCUS AG118925 682 bp DNA linear GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-127D11.F, genomic survey sequence.

ACCESSION AG118925

VERSION AG118925.1 GI:16739444

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.

COMMENT

Sequencing: -21M13
LIBRARYVector : PKS145
R.Site 1 : SacI
R.Site 2 : SacI

Location/Qualifiers

1..682
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-127D11.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES

source

ORIGIN

Query Match 73.6%; Score 36.8; DB 9; Length 682;
Best Local Similarity 85.4%; Pred. NO. 0.0011;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;QY 3 TTAAGCTTATAGCATGCTCCGCTTCCAGTCGGGAAACCTGTGCG 50
|||||
DB 255 TTAATTGTTTGGCGCTCACTGCCGCTTCCAGTCGGGAAACCTGTGCG 302Search completed: May 10, 2005, 07:25:15
Job time : 353.316 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 03:40:34 ; Search time 1072.29 Seconds
(without alignments)

7998.346 Million cell updates/sec

Title: US-09-896-888a-1_COPY_351_527

Perfect score: 177

Sequence: 1 gcttattctgacagcgc.....gcccgcaacgatctggttaa 177

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb_ba:

2: gb_htg:

3: gb_in:

4: gb_om:

5: gb_ov:

6: gb_pat:

7: gb_ph:

8: gb_pl:

9: gb_pr:

10: gb_ro:

11: gb_sts:

12: gb_sy:

13: gb_uni:

14: gb_vi:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	177	100.0	564	BD070856	Insect ex
2	177	100.0	1429	S64501	p8.9=8.9 kd
3	177	100.0	2773	AX766573	Sequence
4	177	100.0	4170	M83827	Orgyia pseu
5	177	100.0	13195	U75930	Orgyia pseu
6	35.2	19.7	129609	AF512031	AF512031 Choriston
7	34.8	19.7	300363	AE016781	AE016781 Pseudomon
8	33	18.6	155274	AC121979	AC121979 Mus muscu
9	32.8	18.5	179571	CR361569	Danio rer
10	32.4	18.3	177263	AP003376	AP003376 Oryza sat
11	32.4	18.3	222256	AC134482	AC134482 Rattus no
12	32.4	18.3	270958	AC135655	AC135655 Rattus no
13	32.2	18.2	299550	AP005031	AP005031 Streptomy
14	32	18.1	921	BD221352	Human gen
15	32	18.1	131158	AY327402	AY327402 Choriston
16	31	17.5	110000	CP000010	Continuation (3 of
17	31	17.5	303450	1 SC0939130	AL939130 Streptomy
18	30.8	17.4	1268	6 CQ597684	CQ597684 Sequence
19	30.8	17.4	1342	6 CQ610737	CQ610737 Sequence

20	30.8	17.4	1426	3	AY070687	AY070687 Drosophil
21	30.8	17.4	3448	6	CQ610736	CQ610736 Sequence
22	30.8	17.4	3600	6	CQ597683	CQ597683 Sequence
23	30.8	17.4	10172	1	AE012382	AE012382 Xanthomon
24	30.8	17.4	17997	2	AC019883	AC019883 Drosophil
25	30.8	17.4	36800	9	HS313D11	Z92544 Human DNA s
26	30.8	17.4	41842	2	AC133915	AC133915 Homo sapi
27	30.8	17.4	93375	3	AC005554	AC005554 Drosophil
28	30.8	17.4	140403	2	AC109444	AC109444 Homo sapi
29	30.8	17.4	164920	2	AC092187	AC092187 Drosophil
30	30.8	17.4	184554	3	AC099022	AC099022 Drosophil
31	30.8	17.4	256073	3	AE006464	AE006464 Homo sapi
32	30.8	17.4	314957	3	AE003581	AE003581 Drosophil
33	30.6	17.3	110000	2	AF331832S1	AF331832 Homo sapi
34	30.6	17.3	125020	9	AF429315_0	AF429315 Homo sapi
35	30.6	17.3	189509	9	AC099564	AC099564 Homo sapi
36	30.6	17.2	41905	9	AC142495	AC142495 Homo sapi
37	30.4	17.2	105642	9	AL732414	AL732414 Human DNA
38	30.4	17.2	119525	9	AC142496	AC142496 Homo sapi
39	30.4	17.2	147127	9	AC142497	AC142497 Homo sapi
40	30.4	17.1	1127	8	AK104298	AK104298 Oryza sat
41	30.2	17.1	1127	8	AK070153	AK070153 Oryza sat
42	30.2	17.1	2021	8	AK072256	AK072256 Oryza sat
43	30.2	17.1	2127	8	AK107045	AK107045 Oryza sat
44	30.2	17.1	33087	2	AC149421	AC149421 Phakopsor
45	30.2	17.1				

ALIGNMENTS

RESULT 1	BD070856	564 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD070856				
DEFINITION	Insect expression vectors.				
ACCESSION	BD070856				
VERSION	BD070856.1 GI:22616459				
KEYWORDS	JP 2001516225-A/1.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 564)				
AUTHORS	Grigliatti,T.A., Theilmann,D.A., Pfeifer,T.A. and Hegedus,D.D.				
TITLE	Insect expression vectors				
JOURNAL	Patent: JP 2001516225-A 1 25-SBP-2001;				
COMMENT	THE UNIVERSITY OF BRITISH COLUMBIA				
	OS Multicapsid nucleopolyhedrovirus				
	PN JP 2001516225-A/1				
	PD 25-SEP-2001				
	PF 26-MAR-1998 JP 1998541010				
	PR 27-MAR-1997 US 60/049946,28-JAN-1998 CA 2221819 PI				
	THOMAS A GRIGLIATTI,DAVE A THEILMANN,THOMAS				
	A PFEIFER,DWAYNE D				
	PI HEGEDUS				
	PC C12N15/85,C12N5/06,C12N15/69//C12N9/22				
	CC Insect expression vectors				
	PH Key				
	FT source				
	FT				
FEATURES	Location/Qualifiers				
source	1..564				
	/organism="unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
ORIGIN					
Query Match	100.0%;	Score 177;	DB 6;	Length 564;	
Best Local Similarity	100.0%;	Pred. No. 6.2e-41;			
Matches 177;	Conservative	0;	Mismatches	0;	Gaps 0;
QY	1	GTCTTATCTGACGAGCCGACGCTTCCTGTTGTCTTAACCGAGCGGACGCACTCCT	60		
DB	351	GTCTTATCTGACGAGCCGACGCTTCCTGTTGTCTTAACCGAGCGGACGCACTCCT	410		

```

QY 61 TATCGGAACAGGACGCGCTCCATATACGCGCGGTTATCTCATGCGGTGACCGGACA 120
DB 411 TATCGGAACAGGACGCGCTCCATATACGCGCGGTTATCTCATGCGGTGACCGGACA 470

QY 121 CGAGGCGCCCTCCGCTTATCGGCTTATAAATACAGCCGCAACGATCTGGTAAA 177
DB 471 CGAGGCGCCCTCCGCTTATCGGCTTATAAATACAGCCGCAACGATCTGGTAAA 527

RESULT 2
S64501/c
LOCUS S64501 1429 bp DNA linear VRL 30-SEP-1993
DEFINITION p8.9=8.9 kda basic protein [Oryza pseudotsugata multicapsid
nuclear polyhedrosis virus OpNPV, Genomic, 1429 nt].
ACCESSION S64501
VERSION S64501.1 GI:404518
KEYWORDS
SOURCE
ORGANISM Oryza pseudotsugata multicapsid nucleopolyhedrovirus
Oryza pseudotsugata multicapsid nucleopolyhedrovirus
Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
Nucleopolyhedrovirus.
REFERENCE 1 (bases 1 to 1429)
AUTHORS Wu, X., Stewart, S. and Theilmann, D. A.
TITLE Characterization of an early gene coding for a highly basic 8.9K
protein from the Oryza pseudotsugata multicapsid nuclear
polyhedrosis virus
JOURNAL J. Gen. Virol. 74 (Pt 8), 1591-1598 (1993)
MEDLINE 93346965
PUBMED 8345350
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 136152] from the original journal article.
FEATURES
source
1..1429
/organism="Oryza pseudotsugata multicapsid
nucleopolyhedrovirus"
/mol_type="genomic DNA"
/db_xref="taxon:164623"
763..990
/gene="p8.9"
/note="8.9 kda basic protein"
763..990
/gene="p8.9"
/note="8.9 kda basic protein"
/codon_start=1
/protein_id="AAB27738.1"
/db_xref="GI:404519"
/translation="MNSWKIRMAKQQQVVRVQRHRAAKRLKLYKAKKLAELCEKIQ
LQRVNDALAKAFEEFVYPHFSPLYTLN"
ORIGIN
Query Match 100.0%; Score 177; DB 14; Length 1429;
Best Local Similarity 100.0%; Pred. No. 6e-41;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCCTTATCGTACAGACGCGCTCCATATACGCGCGGTTATCTCATGCGGTGACCGGACA 120
DB 415 GTCCTTATCGTACAGACGCGCTCCATATACGCGCGGTTATCTCATGCGGTGACCGGACA 356

QY 61 TATCGGAACAGGACGCGCTCCATATACGCGCGGTTATCTCATGCGGTGACCGGACA 120
DB 355 TATCGGAACAGGACGCGCTCCATATACGCGCGGTTATCTCATGCGGTGACCGGACA 296

QY 121 CGAGGCGCCCTCCGCTTATCGGCTTATAAATACAGCCGCAACGATCTGGTAAA 177
DB 295 CGAGGCGCCCTCCGCTTATCGGCTTATAAATACAGCCGCAACGATCTGGTAAA 239

RESULT 3
AX766573
LOCUS AX766573 2773 bp DNA linear PAT 25-JUN-2003
DEFINITION Sequence 60 from Patent WO03042244.
ACCESSION AX766573

```

```

VERSION AX766573.1 GI:32260450
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Klynsen, S., Nielsen, F. S., Bratt, T., Voldborg, B. and Mouritsen, S.
TITLE Novel immunogenic mimetics of multimer proteins
JOURNAL Patent: WO 03042244-A 60 22-MAY-2003;
Pharmexa A/S (DK); Klynsen, Steen (DK); Nielsen, Finn Stausholm
(DK); Bratt, Tomas (DK); Voldborg, Bjorn (DK); Mouritsen, Soren
(DK)
FEATURES
Location/Qualifiers
1..2773
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="p22Op2F expression vector for insect cells"
misc_recomb 561..566
/note="HindIII site"
misc_recomb 573..578
/note="Aval site"
misc_recomb 586..591
/note="EcoRI site"
misc_recomb 593..598
/note="BamHI site"
misc_recomb 625..630
/note="ClaI site"
misc_recomb 629..634
/note="ClaI site"
misc_recomb 1156..1161
/note="ApaLI site"
misc_recomb 2128..2133
/note="PstI site"
misc_recomb 2204..2209
/note="NcoI site"
misc_recomb 2284..2289
/note="Aval site"
misc_recomb 2294..2299
/note="Aval, SmaI, and XmaI site"
misc_recomb 2551..2556
/note="ApaLI site"
ORIGIN
Query Match 100.0%; Score 177; DB 6; Length 2773;
Best Local Similarity 100.0%; Pred. No. 5.9e-41;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCCTTATCGTACAGGACGCGCTTCCTGTTGCTTAACCGGACGCGGACGCACTCCT 60
DB 355 GTCCTTATCGTACAGGACGCGCTTCCTGTTGCTTAACCGGACGCGGACGCACTCCT 414

QY 61 TATCGGAACAGGACGCGCTCCATATACGCGCGGTTATCTCATGCGGTGACCGGACA 120
DB 415 TATCGGAACAGGACGCGCTCCATATACGCGCGGTTATCTCATGCGGTGACCGGACA 474

QY 121 CGAGGCGCCCTCCGCTTATCGGCTTATAAATACAGCCGCAACGATCTGGTAAA 177
DB 475 CGAGGCGCCCTCCGCTTATCGGCTTATAAATACAGCCGCAACGATCTGGTAAA 531

RESULT 4
NPHTTAA
LOCUS NPHTTAA 4170 bp DNA linear VRL 02-AUG-1993
DEFINITION Oryza pseudotsugata nuclear polyhedrosis virus transcriptional
trans-activator (IE-2) gene, complete cds; ORF, complete cds.
ACCESSION M83827
VERSION M83827.1 GI:332540
KEYWORDS
SOURCE Oryza pseudotsugata single capsid nucleopolyhedrovirus
Oryza pseudotsugata single capsid nucleopolyhedrovirus
Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
Nucleopolyhedrovirus.
ORGANISM

```

```

REFERENCE 1 (bases 1 to 4170)
TITLE Theilmann,D.A. and Stewart,S.
Molecular analysis of the trans-activating IE-2 gene of Orgyia
pseudotsugata multicapsid nuclear polyhedrosis virus
JOURNAL Virology 187 (1), 84-96 (1992)
MEDLINE 92142536
PUBMED 1736546
REFERENCE 2 (bases 1 to 4170)
TITLE Theilmann,D.A. and Stewart,S.
Tandemly repeated sequence at the 3' end of the IE-2 gene of the
baculovirus Orgyia pseudotsugata multicapsid nuclear polyhedrosis
virus is an enhancer element
JOURNAL Virology 187 (1), 97-106 (1992)
MEDLINE 92142537
PUBMED 1736547
COMMENT Original source text: Orgyia pseudotsugata nuclear polyhedrosis
virus DNA.
FEATURES
    source
        Location/Qualifiers
            1..4170
                /organism="Orgyia pseudotsugata single capsid
                nucleopolyhedrovirus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10450"
            627..3959
                /gene="IE-2"
            627..634
                /gene="IE-2"
            657..1908
                /gene="IE-2"
            691..1908
                /gene="IE-2"
                /codon_start=1
                /product="transcriptional trans-activator"
                /protein_id="AAA46749.1"
                /db_xref="GI:332541"
                /translation="MRSKNANAPTSPNRRNLSLVGRRLTVSPDPAASAQASPPR
                SAPRAPRVHVGDPGLARASYALPNGVNLHGDAHPPEEDDDILFVDTAAEQ
                RQAVNLHVRNRELGEMTRPTLLNYSPTSPRSPDLIMPDIQ
                AREQPLAFNDSDDRLEQLVMSAEAPQLPAPQVDVSLCHICSTFDIQ
                NYNSVFTSECHNACPKCYVSIIVFGESYKSCINRTTISCRAVRDGYVELSTMS
                TVRDSALIKRHQAQLSDSNMPSNEMTTIQELQALAEALRAATARAHHDVNRASDSQ
                LIRQQLDQKEALAHESNARLKLQKQNETLSANNSLQHQLTQVIESRVKMEQFKRQ
                HDEFMEKPLSLSS"
            1958..1963
                /gene="IE-2"
            2017..2759
                /gene="IE-2"
            2915..2922
                /gene="IE-2"
            2944..3907
                /gene="IE-2"
            2984..3907
                /gene="IE-2"
            /notes="ORF; homologous to the ACNVPV PE-38 gene; putative"
            /codon_start=1
            /protein_id="AAA46750.1"
            /db_xref="GI:332542"
            /translation="MSRRYSPVRIIRNVGRRRVOELLSEFNSRPVTVAANCAVCLE
            TYVCQNNIIDFMPSECTHFCYKCVLNMKNMNVRAAVSCPMCKKVGWQAF
            PNVSCCKIKKTGDTAPCLQPLAKTIQDRYMATEBEAEETFSFVINKLQALDA
            AQKEARDLQESMERQKQAHNVAMNSCEQVTALQTTLADMQALDRSEALSTLAHN
            RAANVQIDSLIRAVQRLAEQAQAPVSNVVEFNDNQNTNLNHERFRSYVYSTVDNMI
            EDSIKLSQSHVFGAACLPCSVNVEINFPFDE"
            3954..3959
                /gene="IE-2"
    polyA_signal
    enhancer
    TATA_signal
    mRNA
    CDS
    polyA_signal
    ORIGIN
    Query Match 100.0%; Score 177; DB 14; Length 4170;
    Best Local Similarity 100.0%; Pred. No. 5.8e-41;
    Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 1 GTCTTATCGTGACAGGACGCCAGCTTCCTGTGTGTTAAACCGCAGCGCACTCCT 60
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||

```

```

Db 480 GTCTTATCGTGACAGGACGCCAGCTTCCTGTGTGTTAAACCGCAGCGCACTCCT 539
QY 61 TATCGGAACAGGACGCGCTCCATATCAGCGCGCGCTTATCTCATCGCGTGAACCGGACA 120
|||||
Db 540 TATCGGAACAGGACGCGCTCCATATCAGCGCGCGCTTATCTCATCGCGTGAACCGGACA 599
|||||
QY 121 CGAGGCGCGCTCCGCTTATCGCGCTATAAATACAGCCGCAACGATCTGGTAAA 177
|||||
Db 600 CGAGGCGCGCTCCGCTTATCGCGCTATAAATACAGCCGCAACGATCTGGTAAA 656

RESULT 5
OPU75930 131995 bp DNA circular VRL 25-MAR-2003
LOCUS Orgyia pseudotsugata multicapsid nucleopolyhedrovirus, complete
DEFINITION genome.
ACCESSION U75930
VERSION U75930.2 GI:11024985
SOURCE
ORGANISM
    Orgyia pseudotsugata multicapsid nucleopolyhedrovirus
    Orgyia pseudotsugata multicapsid nucleopolyhedrovirus
    Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
    Nucleopolyhedrovirus.
REFERENCE 1 (bases 1 to 131995)
AUTHORS Ahrens,C.H., Russell,R.L., Funk,C.J., Evans,J.T., Harwood,S.H. and
Rohrmann,G.F.
TITLE The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear
polyhedrosis virus genome
JOURNAL Virology 229 (2), 381-399 (1997)
MEDLINE 97271300
PUBMED 9126251
REFERENCE 2 (bases 1 to 131995)
AUTHORS Rohrmann,G.F.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1996) Oregon State University, Agricultural
Chemistry, Corvallis, OR 97331-7301, USA
REFERENCE 3 (bases 1 to 131995)
AUTHORS Rohrmann,G.F.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-1998) Oregon State University, Agricultural
Chemistry, Corvallis, OR 97331-7301, USA
REMARK Sequence update by submitter
REFERENCE 4 (bases 1 to 131995)
AUTHORS Rohrmann,G.F.
TITLE Direct Submission
JOURNAL Submitted (26-OCT-2000) Oregon State University, Agricultural
Chemistry, Corvallis, OR 97331-7301, USA
REMARK Sequence update by submitter
COMMENT On Oct 26, 2000 this sequence version replaced gi:2934903.
FEATURES
    source
        1..131995
            /organism="Orgyia pseudotsugata multicapsid
            nucleopolyhedrovirus"
            /mol_type="genomic DNA"
            /db_xref="taxon:164623"
            /note="similar to Autographa californica nuclear
            polyhedrosis virus (AcMNPV) deposited in GenBank Accession
            Number L22858"
            complement(123..947)
            /note="ORF1; pk-1; similar to AcMNPV ORF10"
            /codon_start=1
            /product="protein kinase 1"
            /protein_id="AAC59000.1"
            /db_xref="GI:1911247"
            /translation="MDAALQSLRSFVADCAVLAPKVNGRFGKMDVHLHPTTSKLF
            RKTIAHSFSADEINVDLMSDHPFVDMYFCYSSTAWAIWMDYVPCDLPETLQTLQ
            GADNALVYNIYQLCDALNDLHNATGYIHNDVLEKLVLYFGARDRVLYLCDYGLCKRE
            HSPVHDGTLVEYSPKIRRHNYARSFDWYAGVLYKLTGGRHFFERSADEVLDLAS
            MERRQYNDPAALKNVNRLMARDFFVCLTRFNECRSTDYKQIAKHSPLASRDYI"
            946..2370
            /note="ORF2; similar to AcMNPV ORF9"
            /codon_start=1
            /product="1629-capsid"
    CDS

```

```
/protein_id="AAC59001.1"
/db_xref="GI:2934904"
/translation="MMERQYQSVRSYILINNOHNAIAAGPFLQVAGPEAHSGVGRNVD
RAVLNRQAVLDLKLAEIDYADTAQADPEASRRHFATLNRMLLLIGVQDPDAR
RLNSLRLIALLRLVDVDAENVLSDGFYEVSKYISYQOTFAOTPTPATASQOQT
RLSLRPTSLPQTOASLPQOTPDQPEMSPSPFVHTTALLPQTTQPPATDTSRSP
SDEFVYVKGKRAVPDTRFKPVPKPEHLKSRSSVATNAGATPVAPPPPPPSAD
VTTSMPPPPPPFSGADVTTSMPPPPPPMVDLATSMPPPPPPPMVDLATSMPPIINA
INLLIDAMVAETNKNAGNRSALLDQIKQKTKTKTQPADGAPATPRSTLLSEIRQ
KRLKLKRIEDQSTQTLKDVDTTQTKTKILNFVTNIDRISKQEQEEDRLDTIT
KRRPAVEHTDGNSTGNSDDWRDD"
complement(2533..3270)
CDS
/notes="ORF3; ph; similar to AcMNPV ORF8"
/codon_start=1
/product="polyhedrin"
/protein_id="AAC59002.1"
/db_xref="GI:1911249"
/translation="MPDYSYRTTIGRTVYDKNYKKNLGSVINKAKRKHLEHEDEB
KHLDPHMYAEOPFLGPKNOKLTLEKIRNVKPDTKLIVNMSKGEFLRETWTFR
VEDSPPIVDOEVMDFVLVNMRTPRNRCYKFLAHLRWDCDYVHEVIRIVEPSY
VGMNEYRISLAKGGGCPIMNIIHAETNSFESFVNRVWENFYKPIYIGTDSBEE
EILIEVSLVFKYKEPADAPLETGPAY"
complement(3350..3541)
CDS
/notes="ORF4"
/codon_start=1
/product="unknown"
/protein_id="AAC59003.1"
/db_xref="GI:1911250"
/translation="MRKNNGASQYPEHLTTVHKRHLQLSVCWCDQORSDPHSAKGL
MFEHMFCLRIYVQOKIKH"
complement(3698..4093)
CDS
/notes="ORF5"
/codon_start=1
/product="unknown"
/protein_id="AAC59004.1"
/db_xref="GI:1911251"
/translation="WGASQNSEOKILKLFYWSQTAGALDDEKDLHCLYDLERFVGA
HLNKRADKARKKCAEAKALKRVBIADRHLEAAAPACADDGRNSTLSRAQLDDI
AKEKEIVDRIRHLQKQSLKQDLKKR"
complement(4125..4739)
CDS
/notes="ORF6; left-2; similar to AcMNPV ORF6"
/codon_start=1
/product="late expression factor 2"
/protein_id="AAC59005.1"
/db_xref="GI:1911252"
/translation="WERVWNPAGIDGLKRSETYLVDPHDFVGVLTLSPYTVFERGLF
VMSGRLIALAPKQPPQPAVRFPQRRNRNCLKACADGAQSLAKVLAARVSM
PCWSTMDLSSAPRGNMYRKFEEFNCYLANVITCKTKTACILGALHLYRMDAKV
GEVTHLLIAQDQVYKPSNCAKMKVTKLCPQASMCKGLNPICNF"
complement(4742..4975)
CDS
/notes="ORF7; similar to AcMNPV ORF5"
/codon_start=1
/product="unknown"
/protein_id="AAC59006.1"
/db_xref="GI:1911253"
/translation="WNRPTWRNTAAVTTDYDREQLRRLNSLRVHCLTSATGPD
CNRPLEAGDRAPAVIKAAAGQSHSLCDKV"
5020..5460
CDS
/notes="ORF8; similar to AcMNPV ORF4"
/codon_start=1
/product="unknown"
/protein_id="AAC59007.1"
/db_xref="GI:1911254"
/translation="MPGARFVPSRLTQEFKENVVAVHLMGLRALIDGKVTISADY
RRFGFLSRNLVSAMVNVQYAPDATIDMRHQTITFRVCONCHADVPAPDDHS
IARYLLAECGAVLVIDHLDVFGTEEGVNELLEVORINAGGDL"
complement(5669..6151)
CDS
/notes="ORF9; ptp-2"
/codon_start=1
/product="protein tyrosine phosphatase 2"
/protein_id="AAC59008.1"
/db_xref="GI:1911255"
/translation="MYEASQIDHVFVGGYGDNEAMLRFIENHAIASVISLIDSDVA
```

```
PIRHALGLPVGDHIVHYCBAAPTCAALPNAMPALYDMVVRRIIGEGKRVLVHCYAGASR
SAALAYYLMSRSPQMPYQALNAVQSKRVAINDHFVFLATCSYRFVNDLKPQIV
"
complement(6129..6791)
CDS
/notes="ORF10; ptp-1; similar to AcMNPV ORF1"
/codon_start=1
/product="protein tyrosine phosphatase 1"
/protein_id="AAC59009.1"
/db_xref="GI:1911256"
/translation="MLTSRRWAVIYTRTSERNCGAWCTNGVRRRRQVHLPLSVARAM
PDRMEHTACGAVIEGRLLCFKFNALBELFVTSDEDRWTAASVLAHLSALGAVI
DLNTARYYDGAQWKMGLYKKIRVGRGAVPDDDIVASFIETVDFEFRCPMLNAV
HWTGLNRSGLVCRVWVERLGSPTDAITARPETARGHKIERTNYLQDLLARKHVRGQ
PN"
6819..7814
CDS
/notes="ORF11; similar to AcMNPV ORF11"
/codon_start=1
/product="unknown"
/protein_id="AAC59010.1"
/db_xref="GI:1911257"
/translation="MSLSKLLVYVYGSYNLPHDRYGSYHLVRIVHEHLTNTYVSN
ASCVRDIATARCINSGHLCFDVARQLDVSVAARLSAWFRGCDATGLCADQRALA
DIDRHAPLARRVGRANI FALDAIADIPSDVTNNLQIGRFMHFPCSGIARVAVF
DPDIRADGWMYHKFCVLTVMHLVACGAVPAGSATRLERDAVAKHGNDEGNCAPATAA
VYGRFCAIGRHEFAHKTACMHLILFQFMRNDLTPADERHFCFVGIVKDFGRCKDTYD
LRTHADALYHGTTRQKNAFLDLLCCVNASDIDADCYDCVNVKFTATQNKYKM"
complement(8180..9142)
CDS
/notes="ORF12; similar to AcMNPV ORF13"
/codon_start=1
/product="unknown"
/protein_id="AAC59011.1"
/db_xref="GI:1911258"
/translation="MFSWMEGWNAADEOVNAEFDQAYRYAVDORAHSDLVRDVF
RCHPFPKFRYVILDETAGRCSSVDFCKGLKISHDLLQRCNEDRQVROLNELVLGAP
PAQPDLSGLSFATKGLVOLLOQPSFANKNEVLVANGANDHNRDLNLLKIEAVLNHV
KTINTNSDFEISAKHSFKLEVGRFQORLOTLTKNALCAAPRTAPGVVFR
DVTXKPHLAVFMVEDRGVQTAFARGQEFKRFKLEFEFGMDVDRGRAPNPLLA
VHCIEFANGHGKIRLPKPKVIEVDCAVNVAKDIVKALINKT"
complement(9094..9825)
CDS
/notes="ORF13; left-1; similar to AcMNPV ORF14"
/codon_start=1
/product="late expression factor 1"
/protein_id="AAC59012.1"
/db_xref="GI:1911259"
/translation="MAPCKYTPERVOMMDAIAYNDSRRILAFMTDRPRWVHAHNFDS
AAQLPVIYVNSISDVHKPLESGREWIDADFCADCAELMLKVNGATVAMLPFF
EGKEDAVQRMFSGNRGFHLWLPCKEKFMDAPKSLREHNVFKOPAKLYSGDIRPG
SFADCVRRVAVHYIGDAREDLVRYWPDVDVFCNANKQIRAPFSYNYKGGDYSRCL
TQQLQORIKACAGCLAGGTPPTSK"
9893..11362
CDS
/notes="ORF14; egt; similar to AcMNPV ORF15"
/codon_start=1
/product="ecdysteroid UDP-glucosyltransferase"
/protein_id="AAC59013.1"
/db_xref="GI:1911260"
/translation="MVFLIILTLATGARAASILAVPTPASHVHVYRAYVHALVK
NCHNVYIKOLLDYAVQDECGRVFLTSNRTFDVAVVEAFADYALVFGHLFRPAPVIOI
TADNVYGLIEMFKDQPDNVRFRVLEQDADMSAQYKKLVASSGVFRKRGVADETTV
APGVGLAENFERRAVARHPLHYETFGAALTRGGALSEWRLNEFELLARRSDELL
KQPKGKSTPIIRQLRDNVQLLLNLHPVYDNNRPVPSVQYLGGLHLAQLPORLDA
PLERLINESVDGAVYVSGGIDTNSIHAFLQMLLDTFANNNYTVLWKVDVAAS
VALPRNLVLAQKFSQTAVLRKHNVAFTVQAGLSDEALQARVPMVCLPMWGQOFHH
ARKLQPVQAGRALDTAAVASPQLQLAIREVIADGEAYRARIDKLRAVVEHDAAPDEKA
VKFERVINKNDVNPARSLLKTTAANNAYSDFVRFPL"
11547..12140
CDS
/notes="ORF15; similar to AcMNPV ORF16"
/codon_start=1
```

Query Match 100.0%; Score 177; DB 14; Length 131995;
Best Local Similarity 100.0%; Pred. No. 5.2e-41;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCTTATCGTGACAGGAGCCAGCTTCTGTGTGCTTAACCGAGCCGAGCGACACTCTCT 60

Db 128522 GTCTTATCGTGACAGGACGCCAGCTTCCTGTGTCTTAACCGCGCGACGCAACTCCT 128581
 Qy 61 TATCGGAACAGGACGCCCTCCATATCAGCGCGCGTTCATCTCATGCGCGTGCACCGGACA 120
 Db 128582 TATCGGAACAGGACGCCCTCCATATCAGCGCGCGTTCATCTCATGCGCGTGCACCGGACA 128641
 Qy 121 CGAGGCGCCGTCGCGCTTATCGCGCTATAAATACAGCCGCGCAACGATCTGGTAAA 177
 Db 128642 CGAGGCGCCGTCGCGCTTATCGCGCTATAAATACAGCCGCGCAACGATCTGGTAAA 128698
 RESULT 6
 AF512031/c 129609 bp DNA circular VRL 27-MAY-2004
 LOCUS
 DEFINITION Choristoneura fumiferana MNPV polyhedrin, complete genome.
 AF512031 AF177329 S78506 S81690 U10441 U18677 U26676 U26734 U53854
 ACCESSION
 U57401 U59008 U70432 U72240 X65395 S46001
 VERSION AF512031.2 GI:47157118
 KEYWORDS
 SOURCE Choristoneura fumiferana MNPV
 ORGANISM Choristoneura fumiferana MNPV
 Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 Nucleopolyhedrovirus.
 REFERENCE 1 (bases 1 to 129609)
 AUTHORS Lee,H.Y., Arif,B., Dobos,P. and Krell,P.
 TITLE Identification of bent DNA and ARS fragments in the genome of
 Choristoneura fumiferana nuclear polyhedrosis virus
 JOURNAL Virus Res. 24 (3), 249-264 (1992)
 MEDLINE 93033705
 PUBMED 1413988
 REFERENCE 2 (bases 1 to 129609)
 AUTHORS Xie,W.D., Arif,B., Dobos,P. and Krell,P.J.
 TITLE Identification and analysis of a putative origin of DNA replication
 in the Choristoneura fumiferana multinucleocapsid nuclear
 polyhedrosis virus genome
 JOURNAL Virology 209 (2), 409-419 (1995)
 MEDLINE 95297142
 PUBMED 7778276
 REFERENCE 3 (bases 1 to 129609)
 AUTHORS Liu,J.J. and Carstens,E.B.
 TITLE Identification, localization, transcription, and sequence analysis
 of the Choristoneura fumiferana nuclear polyhedrosis virus DNA
 polymerase gene
 JOURNAL Virology 209 (2), 538-549 (1995)
 MEDLINE 95297155
 PUBMED 7778286
 REFERENCE 4 (bases 1 to 129609)
 AUTHORS Barrett,J.W., Krell,P.J. and Arif,B.M.
 TITLE Characterization, sequencing and phylogeny of the ecdysteroid
 UDP-glucosyltransferase gene from two distinct nuclear polyhedrosis
 viruses isolated from Choristoneura fumiferana
 J. Gen. Virol. 76 (Pt 10), 2447-2456 (1995)
 JOURNAL 96030854
 MEDLINE
 PUBMED 7595348
 REFERENCE 5 (bases 1 to 129609)
 AUTHORS Giu,W., Liu,J.J. and Carstens,E.B.
 TITLE Studies of Choristoneura fumiferana nuclear polyhedrosis virus gene
 expression in insect cells
 JOURNAL Virology 217 (2), 564-572 (1996)
 MEDLINE 96183379
 PUBMED 8610448
 REFERENCE 6 (bases 1 to 129609)
 AUTHORS Liu,J.J. and Carstens,E.B.
 TITLE Identification, molecular cloning, and transcription analysis of
 the Choristoneura fumiferana nuclear polyhedrosis virus
 spindle-like protein gene
 JOURNAL Virology 223 (2), 396-400 (1996)
 MEDLINE 96400202
 PUBMED 8806578
 REFERENCE 7 (bases 1 to 129609)
 AUTHORS Lapointe,R., Back,D.W., Ding,Q. and Carstens,E.B.
 TITLE Identification and molecular characterization of the Choristoneura

fumiferana multicapsid nucleopolyhedrovirus genomic region encoding
 the regulatory genes pkip, p47, lef-12, and gta
 JOURNAL Virology 271 (1), 109-121 (2000)
 MEDLINE 20276145
 PUBMED 10814576
 REFERENCE 8 (bases 1 to 129609)
 AUTHORS Carstens,E.B., Liu,J.J. and Dominy,C.
 TITLE Identification and molecular characterization of the baculovirus
 CfMNPV early genes: ie-1, ie-2 and pe38
 JOURNAL Virus Res. 83 (1-2), 13-30 (2002)
 MEDLINE 21854555
 PUBMED 11864738
 REFERENCE 9 (bases 1 to 129609)
 AUTHORS de Jong,J.G., Dominy,C.N., Lauzon,H.A., Arif,B.M., Carstens,E.B.
 and Krell,P.J.
 TITLE Complete Genome of Choristoneura fumiferana Multiple
 Nucleopolyhedrovirus
 JOURNAL Unpublished
 REFERENCE 10 (bases 1 to 129609)
 AUTHORS de Jong,J.G., Dominy,C.N., Lauzon,H.A., Arif,B.M., Carstens,E.B.
 and Krell,P.J.
 TITLE Direct Submission
 JOURNAL Submitted (13-MAY-2002) Department of Microbiology, University of
 Guelph, Chemistry and Microbiology Building, Guelph, ON N1G 2W1,
 Canada
 REFERENCE 11 (bases 1 to 129609)
 AUTHORS de Jong,J.G., Dominy,C.N., Lauzon,H.A., Arif,B.M., Carstens,E.B.
 and Krell,P.J.
 TITLE Direct Submission
 JOURNAL Submitted (13-MAY-2004) Department of Microbiology, University of
 Guelph, Chemistry and Microbiology Building, Guelph, ON N1G 2W1,
 Canada
 REMARK
 COMMENT Sequence update by submitter
 FEATURES On May 13, 2004 this sequence version replaced gi:30269978.
 source
 1. 129609
 /organism="Choristoneura fumiferana MNPV"
 /mol_type="genomic DNA"
 /db_xref="taxon:208973"
 /country="Ireland"
 complement(1..735)
 /note="ORF1; Ac8/Op3"
 /codon_start=1
 /product="polyhedrin"
 /protein_id="AAP29795.1"
 /db_xref="GI:30269979"
 /translation="MPDYSYRPTIGRTYVYDNKYKLGSVIKRKKKHLLEHDEEK
 HLDPLDHVNAEDPFLGPGKQKLTFLKREINVKPDTMKLIWNVSGKEFLRETWTRFV
 EDSFTVNDQVMDVFLVYNNRPTPNRCYKFLRQHALRWDGCDYVPHEVIRIYEPSTV
 GNNYRISLAKKGGGCGPIMTHAEYTNFSFESFVNRVWENFYKPIVYIGTDSGEBEE
 MLIETSLVPKVEFADAPLFTGPAY"
 complement(838..1449)
 /note="ORF2; Op5"
 /codon_start=1
 /product="unknown"
 /protein_id="AAP29796.1"
 /db_xref="GI:30269980"
 /translation="MQGVDFRRAALLDQCGMGRGDAFVDSBEHTIOAVOLRENE
 GRKVVPSKRVQGLSNLQLLVGVSGHGNMAASDKILELFRWSSQTGCEMLDDEK
 GHWKLYDLERFVGAHLNRKCDARAKNKAERAAKRVQISADRHAALEAAAAPIDADD
 QMLKSLTSAQLDEIAREKDIVDRIVQLQKLDRLIKMDGLKKQ"
 complement(1251..1859)
 /note="ORF3; lef2; Ac6/Op6"
 /codon_start=1
 /product="late expression factor 2"
 /protein_id="AAP29797.1"
 /db_xref="GI:30269981"
 /translation="MDQVNPAGAGSVKKPETYLIDPNDPFGVLMLELTPYTVPERGLF
 IRMSGMRLLALLAAPKPKVPTTTRRFPQSRKNVCLKECADGQSLTKVLITARNMPLC
 MSKMADLGSAPRGNNYKRKRFENCYLANVLCTCKCAACLIGALLHFRMDACKVGE
 VTKHLIKAENTYKPSNCARAKAVTKLCPKANNCKGLNPICNY"
 complement(1862..2092)
 /note="ORF4; Ac5/Op7"

```
/codon_start=1
/product="unknown"
/protein_id="AAP29798.1"
/db_xref="GI:30269982"
/translation="MTRPTWNAANAADYDREQLRRDLNLSRSLVHLCSTRSTGTGFD
CNRFUDGVDKAPAVIIKPAAGQHSLLCDKV"
CDS
2137..2568
/note="ORF5; Ac4/Op8"
/codon_start=1
/product="unknown"
/protein_id="AAP29799.1"
/db_xref="GI:30269983"
/translation="MPNFKPSLRTRESKENIVAHIGLSRVRELIDRNVTPADVR
FRFPDRALLAACMVAVQAGDGTIFLOPTLIYVCRNRAVADVPAPDHFIAR
YLMFECGTPLVIDRHLVFGDTGEVNELLEVQRINAGDDL"
CDS
2669..2860
/note="ORF6; Ac2"
/codon_start=1
/product="baculovirus repeated ORF"
/protein_id="AAP29800.1"
/db_xref="GI:30269984"
/translation="MDRNKKNAVTDTWIFKNCIHYTRTGIIHPCLSNFSVRRPVYLO
PHTVLIINNYAQDAPOFAI"
complement(2896..3699)
/note="ORF7; Ac136/Op132"
/codon_start=1
/product="p26"
/protein_id="AAP29801.1"
/db_xref="GI:30269985"
/translation="MAMLTFLLCVSSAALSVMKVTINNVLTYTNDTKTIAVKQVD
DKPFIQVTPQSFTKQELDLHHPGVSANWFPRIANNTKLTVLLSDGSLATIT
DVRVTVPHSHKRMVYQGLYSFALSNSFLANQIYGAPIFEKRMVSVITAREHDYK
NKLVIYPTGISARGLVSGQINFDLIQTLKLGSSVYQKQPLKALKDYALSTNR
NNKPLGPRNAVYFNEDRITIALVEGEFIDRLRSLGPDILRNKQQ"
/note="ORF8; ptp2; Op9"
/codon_start=1
/product="protein tyrosine phosphatase 2"
/protein_id="AAP29802.1"
/db_xref="GI:30269986"
/translation="MYDANQIDKIFYVGGYGGDDKAMLOFIKKYDIASVISLINADVG
PIRQALGPAGEIHVYCEDAPTVALPENAFALYEMTRRIGEGKRVLIHCYAGER
SAALVYVYMRSRQSYEEALSILVKNKRRVAISNHFVRLASKGCKYFVNNVLKIRVS
"
complement(4198..4731)
/note="ORF9; pcp1; Ac1/Op10"
/codon_start=1
/product="protein tyrosine phosphatase 1"
/protein_id="AAP29803.1"
/db_xref="GI:30269987"
/translation="WEPDRMHEFTPCGRVIDGTRLICFKVPLSAELPEYVNTDEDRTW
VASLTRSALGAVIDLITNTRYDGEQIMIREGLLYKKIRVPGAIPEDTVQKFSA
VDEFQDRCPFTMLVGVHCTHGLNRSGLVCRYNVDKLGVSADAIIRFEARGHKIERA
NYLDLLARNHVRREP"
CDS
4800..5819
/note="ORF10; Ac11/Op11"
/codon_start=1
/product="unknown"
/protein_id="AAP29804.1"
/db_xref="GI:30269988"
/translation="MHSVRLNFDNSNLSKLLVYAYCAFNLHLEKYESYHLYRIY
HEHLTETVSNACSVRRDIATACRCFENGFCNRLARQLLDVTDVAARLAAYNRGDKTG
LCANVQLAAETDKYAPLEKRVSGIGNIPALDTIADIPNALDDPQTIIYSGKDFVD
MNNLAHADVDFDPDKIKAEGWYTKFCVLITMRLTVAVPTELMTLQDAVLFKVQ
PQKNGCAPAMAVYRGCGIKRHFQSKAASWILFYQMRNNITPKDERHPSFGVI
KDFGLRCKETITDLRAEADLLYNATIDKKNALFDLLCCVNAADLVDYDIYVDNF
YFICK"
complement(5854..6387)
```

Query Match 19.9%; Score 35.2; DB 14; Length 129609;
Best Local Similarity 92.5%; Pred. No. 11;
Matches 37; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
137 CTTATCGCGCTATAATACAGCCGCAACGACTTCGTAA 176
|||||
Db 124634 CTTATCGCGCTATAATACAGCCGCAACGACTTCGTAA 124595
|||||

RESULT 7
LOCUS 300363 bp DNA linear BCT 12-DEC-2002
DEFINITION Pseudomonas putida KT2440 section 8 of 21 of the complete genome.
ACCESSION AE016781 AE015451
VERSION AE016781.1 GI:26557025
KEYWORDS
SOURCE Pseudomonas putida KT2440
ORGANISM Pseudomonas putida KT2440
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 300363)
Nelson,K., Paulsen,I., Weinel,C., Dodson,R., Hilbert,H., Fouts,D.,
Gill,S., Pop,M., Martins Dos Santos,V., Holmes,M., Brinkac,L.,
Beanan,M., DeBoy,R., Daugherty,S., Kolonay,J., Madupu,R.,
Nelson,W., White,O., Peterson,J., Khouri,H., Hance,I., Lee,P.,
Holtzapple,E., Scanlan,D., Tran,K., Moazzaz,A., Utterback,T.,
Rizzo,M., Lee,K., Kosack,D., Moestl,D., Wedler,H., Lauber,J.,
Hoheisel,J., Straetz,M., Helm,S., Kiewitz,C., Eisen,J., Tilmis,K.,
Duesterhoft,A., Tummler,B. and Fraser,C.
Complete genome sequence and comparative analysis of the
metabolically versatile Pseudomonas putida KT2440
Environ. Microbiol. 4 (12), 799-808 (2002)
2 (bases 1 to 300363)
Nelson,K., Paulsen,I., Weinel,C., Dodson,R., Hilbert,H., Fouts,D.,
Gill,S., Pop,M., Martins Dos Santos,V., Holmes,M., Brinkac,L.,
Beanan,M., DeBoy,R., Daugherty,S., Kolonay,J., Madupu,R.,
Nelson,W., White,O., Peterson,J., Khouri,H., Hance,I., Lee,P.,
Holtzapple,E., Scanlan,D., Tran,K., Moazzaz,A., Utterback,T.,
Rizzo,M., Lee,K., Kosack,D., Moestl,D., Wedler,H., Lauber,J.,
Hoheisel,J., Straetz,M., Helm,S., Kiewitz,C., Eisen,J., Tilmis,K.,
Duesterhoft,A., Tummler,B. and Fraser,C.
Direct Submission
Submitted (05-NOV-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
1..300363
/organism="Pseudomonas putida KT2440"
/mol_type="genomic DNA"
/strain="KT2440"
/db_xref="taxon:160488"
78..1985
/locus_tag="Pp1883"
/locus_tag="Pp1883"
/locus_tag="Pp1883"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAN67502.1"
/db_xref="GI:24983391"
/translation="MLPTISFARGSGAVLYLCOLSNRKCARRYSYMDPVDVSHLN
ERNVSDLDALGSADMTWISYIGIAEHLIYINRWCKADNSAEDLFGEREVLSDLA
GRGELIDNALVYALRGQVYFSTRLDNLNAGNVPQESKRRFFFDVDPAGIWLPAVA
HQDAHGVIDVNLINPEGILLATAPAAKADKDFLSWEPWLNESLPGVPFQVLE
VQESDQGRLLWHLDIGQVFMVGGFAPLHYRIEADGTSSRSPVQDFDIYVPAAGDP
PAPLLAAPRIPGATDGLDPDDTYRAGLVVDYVPELGLSDSLVLYAVGPVTLR
SLRVDSASVDSKRLTFLHREWLQSEVNRGKRFAPSYEFQGRQSRQSVLDLVLQKR
LYLPLTIIPATQSDGASVQGVHPHRSLESGAKVIRIPAEALGDGAPAPTVNLHWE
HGSDGAIVPRPSPGSTRLTFTVPSAMAANLCRLWYTYTTRAGTPTSTSEFLR
VQSTFQDSQPIQIYQEGENRELKLEHVPVGGALCTLPFWTFMAGQLLGTEALGRSAS
GQDVTILRPPTTEAVTYDEYSDGVKAYLPKQFLGNLDLGLDFRLTVSASFDDGETWR
AFPIVDLKLVS"
2024..4489
/locus_tag="Pp1884"
/locus_tag="Pp1884"
/locus_tag="Pp1884"
/note="identified by Glimmer2; putative"
```


QY 121 CGAGGC 126

Db 160936 CAAGGC 160931

RESULT 8

AC121979

LOCUS AC121979 155274 bp DNA linear ROD 13-NOV-2003

DEFINITION Mus musculus BAC clone RP24-289L14 from chromosome 9, complete sequence.

ACCESSION AC121979

VERSION AC121979.3 GI:24476151

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 155274)

AUTHORS McEellan, M., Cotton, M., Doebber, A. and Schatzkamer, K.

JOURNAL The sequence of Mus musculus BAC clone RP24-289L14

REFERENCE 2 (bases 1 to 155274)

AUTHORS Wilson, R.

TITLE Sequencing of Mus musculus

JOURNAL Unpublished (2001)

REFERENCE 3 (bases 1 to 155274)

AUTHORS McPherson, J.D. and Waterston, R.H.

JOURNAL Direct Submission

REFERENCE 4 (bases 1 to 155274)

AUTHORS McPherson, J.D. and Waterston, R.H.

JOURNAL Direct Submission

REFERENCE 5 (bases 1 to 155274)

AUTHORS McPherson, J.D. and Waterston, R.H.

JOURNAL Direct Submission

REFERENCE 6 (bases 1 to 155274)

AUTHORS Wilson, R.

JOURNAL Direct Submission

COMMENT Submitted (13-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Nov 2, 2002 this sequence version replaced gi:22476161.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu>

Contact: submissions@watson.wustl.edu

----- Summary Statistics

Center project name: M_BB0289L14

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see

http://genome.wustl.edu

SOURCE INFORMATION:
The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.

FEATURES

source

1..155274

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="9"

/map="9"

/clones="RP24-289L14"

/clone_lib="RPCI-24"

1..165

/rpt_family="B2"

178..287

/rpt_family="Alu"

288..309

/rpt_family="B4"

437..551

/rpt_family="B2"

568..732

/rpt_family="B2"

788..936

/rpt_family="B4"

1010..1081

/rpt_family="WIR"

1104..1269

/rpt_family="MaLR"

1130..1302

/rpt_family="Alu"

4855..4982

/rpt_family="Alu"

4876..4987

/rpt_family="B4"

5144..5254

/rpt_family="Alu"

5316..5565

/rpt_family="B4"

5566..5711

/rpt_family="Alu"

5712..5786

/rpt_family="B4"

5788..5975

/rpt_family="B2"

5793..5865

/product="tRNA-Cys"

/note="Likely pseudogene (HMM Sc=39.91 / Sec struct Sc=-11.22)"

7045..7194

/rpt_family="Alu"

7670..7741

/rpt_family="Alu"

8319..8498

/rpt_family="7SLRNA"

8388..8550

/rpt_family="B4"

8826..9027

/rpt_family="B2"

8832..8904

/product="tRNA-Ser"

/note="Likely pseudogene (HMM Sc=33.79 / Sec struct Sc=-6.89)"

9947..10098

/rpt_family="B4"

10110..10490

/rpt_family="MaLR"

```
repeat_region 11038..11227
/rpt_family="B2"
tRNA 11043..11114
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=33.19 / Sec struct
Sc=-10.52)"
repeat_region 11238..11348
/rpt_family="B2"
repeat_region 11348..11466
/rpt_family="B2"
repeat_region 11349..11370
/rpt_family="B4"
repeat_region 11472..11624
/rpt_family="RMER6B"
repeat_region 11625..11800
/rpt_family="B2"
repeat_region 11810..11946
/rpt_family="Alu"
repeat_region 11959..12092
/rpt_family="RMER6B"
repeat_region 12093..12290
/rpt_family="B2"
repeat_region 12291..12465
/rpt_family="RMER6B"
repeat_region 12916..13104
/rpt_family="B2"
repeat_region 13112..13274
/rpt_family="ERVK"
repeat_region 14294..14440
/rpt_family="Alu"
repeat_region 14507..14582
/rpt_family="Alu"
repeat_region 15774..16035
/rpt_family="RMER6B"
repeat_region 16036..16182
/rpt_family="Alu"
repeat_region 16183..16352
/rpt_family="RMER6B"
repeat_region 16353..16501
/rpt_family="Alu"
repeat_region 16502..16576
/rpt_family="RMER6B"
repeat_region 17406..17489
/rpt_family="L2"
repeat_region 17585..17934
/rpt_family="ERVL"
repeat_region 17672..17818
/rpt_family="Alu"
repeat_region 17819..18242
/rpt_family="MaLR"
repeat_region 18243..18309
/rpt_family="L2"
repeat_region 18767..18948
/rpt_family="B2"
repeat_region 18962..19079
/rpt_family="B4"
repeat_region 19415..19480
/rpt_family="B2"
tRNA 19420..19492
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=20.93 / Sec struct
Sc=-0.89)"
Query Match 18.6%; Score 33; DB 10; Length 155274;
Best Local Similarity 51.7%; Pred.No. 48;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
Qy 3 CTTATCGTCAGGACGCCAGCTTCCTGTTGCTAACCGCAGCGGACCACTCCTTA 62
|||||
Db 45418 CTTAATGTATCCCGCCCGCTTCCTTTCTCCGTAGCTGACGCGGACAGCGGCTTT 45477
|||||
Qy 63 TCGGAACAGGACGCGCTCCCATATCAGCCGCGGCTTATCTATCGCGGTGACCGGACAG 122
|||||
```

```
Db 45478 CTTACACAGCTCGCACTCTCTCGGGGGCGACGGCTCGCGGTGCCCTGCCACACAGG 45537
Qy 123 AGGCGCCCGTCCCGCTTATCGGCC 147
Db 45538 GACAGAGCGTCCCGCTTCCTCTCC 45562
RESULT 9
LOCUS CR361569/c 179571 bp DNA linear HTG 11-OCT-2004
DEFINITION Danio rerio clone CH211-99120, WORKING DRAFT SEQUENCE, 12 unordered
pieces.
ACCESSION CR361569
VERSION CR361569.5 GI:54035037
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 179571)
McLaren,S.
Direct Submission
Submitted (10-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Oct 11, 2004 this sequence version replaced gi:51592229.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: ZC99120
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 175712 bases at least Q40
Consensus quality: 176324 bases at least Q30
Consensus quality: 176923 bases at least Q20
Insert size: 178471; sum-of-contigs
Insert size: 179905; 5.5% error; agarose-fp
Quality coverage: 7.91x in Q20 bases; sum-of-contigs Quality
coverage: 8.06x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 18211: contig of 18211 bp in length
* 18212 18311: gap of 100 bp
* 18312 33141: contig of 14830 bp in length
* 33142 33241: gap of 100 bp
* 33242 37691: contig of 4450 bp in length
* 37692 37791: gap of 100 bp
* 37792 44623: contig of 6832 bp in length
* 44624 44723: gap of 100 bp
* 44724 90384: contig of 45661 bp in length
* 90385 90484: gap of 100 bp
* 90485 98331: contig of 7847 bp in length
* 98332 98431: gap of 100 bp
* 98432 102255: contig of 3824 bp in length
* 102256 102355: gap of 100 bp
* 102356 111881: contig of 9526 bp in length
* 111882 111981: gap of 100 bp
* 111982 114150: contig of 2169 bp in length
* 114151 114250: gap of 100 bp
* 114251 134658: contig of 20408 bp in length
* 134659 134758: gap of 100 bp
```


3405. 3547,5098. 5199,6055. 6167,6605. 6690,7352. 74911)
/gene="OSUNBa0014K08.1"
/note="hypothetical protein"
/codon start=1
/protein id="BAC05577.1"
/db_xref="GI:21902028"
/translation="MAEPDWILRRERQVLELDMEELOVEVDVAGSSSESDVT
FLNTHDGGSRSEALAFNTSVSLPTCDGVVLFEALPVRVQPSRLTAVDKAV
NHYDAPCMIVHIIHQLDGSSNVTRGQNRFLRRRWTDADVALISRSDGPVGA
YVQFGSCQEMTVYATGLALRCAPSKAHSLPFGTATPQRRGFKKRCHHGHDQD
ATAHLKRCPPGLEKLAQKNGSPORRGAPRLRVPTGNEEDDDPPLAPALPLGR
OAMCCGCPAPAGARVRPRLTPRRHGPCATTPSRHGLDDPPHGLAGLQAGAE
RRPPRPGRPRRAAIRLRLPRLTNLYAEPQLSASAVPGRSPRPPRRRLTAG
RHDRMPSRHLRPSRAAQIQAREGRIRPWGRLIRRLPLTLP"
join(8072. 8086,9178. 9387,9474. 9671,12990. 13120,
14072. 14144,14960. 15376)
/gene="OSUNBa0014K08.2"
join(8072. 8086,9178. 9387,9474. 9671,12990. 13120,
14072. 14144,14960. 15376)
/gene="OSUNBa0014K08.2"
/note="hypothetical protein"
/codon start=1
/protein id="BAC05578.1"
/db_xref="GI:21902029"
/translation="MRSMFLVPLRLACGTGTTPVPPPTASLRRELASPNLPARGG
VACQPNATREAPAAVSSARVYVSPSLARRRTPPPPPPLFPSPFSAASP
FGIIPMRSLRLSGRRRGESGDQAGALILSWVDANYGLDEQFQVEDLQDEASP
NFVEVVTLPFKDGRLMNIEQALCWRSPQKREHAVLRKLESYFOLCSTWAE
QKGLADRLAKPSQYSMEKSATNRAIPDFRSAAAMSDAMRRARRSPCHARNR
NGTRAGARDPPSPSQPIKVLGRLSDVETWVTPAAAVFLRHGETRAPALGTA
GRRRGDDMSVIR"
join(16191. 16218,17138. 17280)
/gene="OSUNBa0014K08.3"
join(16191. 16218,17138. 17280)
/gene="OSUNBa0014K08.3"
/note="hypothetical protein"
/codon start=1
/protein id="BAC05579.1"
/db_xref="GI:21902030"
/translation="MAITVENQPPSVGHCCIMRPSDVAAATCRVQACSCITAAPGGT
GLLRNAGLSSR"
join(19222. 19261,19694. 19848,19915. 20022,20628. 20832,
21203. 21300,21397. 21480,21572. 21649,21807. 21899,
21991. 22020,22277. 22421,22872. 22967,23073. 23252,
23347. 23453,23552. 23653,23956. 24011,24094. 24347,
24422. 24585,24909. 25032,25170. 25423)
/gene="OSUNBa0014K08.4"
join(19222. 19261,19694. 19848,19915. 20022,20628. 20832,
21203. 21300,21397. 21480,21572. 21649,21807. 21899,
21991. 22020,22277. 22421,22872. 22967,23073. 23252,
23347. 23453,23552. 23653,23956. 24011,24094. 24347,
24422. 24585,24909. 25032,25170. 25423)
/gene="OSUNBa0014K08.4"
/note="contains ESTs AU065754 (S10005), AU175140 (S10005)"
/codon start=1
/product="putative kinesin-like protein"
/protein id="BAC05580.1"
/db_xref="GI:21902031"

/translation="MGTVNGEYDFDAANRAEVDWLGGLPEFPLDLSDEELRD
YLINGCALVADMPGLVLEGTVGGYASDQSRNVKFLSVVAEMGLPGFGVKDLEEG
SMSSIVECILAADKONVATOLGGHISNSTAKPIRRKLRLRDEDPVLSVATPKYPK
SQORSPLSGKINVEVQFKGTYTDLPAKISEMLHNSLDNAPQCSLLRVVNGILD
ESTLRKGEPIPHRVHLNRVQIEIHRIGIQADHTRNQNSIINKTREDKYSKIALE
TLVNGNEEMAINLEVVKESKIDERRKLGQEDMIRLIREKENAENIATSLHQE
MQVMNRHGFQDFRGLVDAENYHKLVAENQKLFNEQELKGNIRVYCRVPFLPG
QDGLTAIDYIGNGHILLANPSKQKGYRMFKNFVQTHSSQAEVFSDIQPLIRS
VLDFGNVICPAYQGTGKTYTWSGPTREDMGVNRALNDLFDI SLRKNAFSVES
NRVAGTALNERSRSRSHLTVHVRGLDVNGKSTSRGCLHLIDLASQVSESEATGD
RLKEAQRHINKSLAGDVIFSLQAQNAHPYRNKLTQVLQSSLGQAKLMPFVQNG
DIESYSETISLFABRVSGVEIGAARNSREGKDIKELLEQVASLKDTIARKDMEIEQ
LQLKSKSPNSMTRNGSLLRQSTSTGLSLPVSASQOQLSVVTLPLADTGLAG
SVFEAEADNADGCGSVGETEYPAGASSETSAERAVYLILFSPISCAKPSCACTSGE
LMIVPVRLRTGT"

join(27978. 28537,28795. 29152,29378. 29650)
/gene="OSUNBa0014K08.5"
join(27978. 28537,28795. 29152,29378. 29650)
/gene="OSUNBa0014K08.5"
/note="hypothetical protein"
/codon start=1
/protein id="BAC05581.1"
/db_xref="GI:21902032"
/translation="MATASVTVTLQRPRLPVPPEFRAPPSPVTGLTSSSSGLAYG
ECPDSNEDDEIGRFLRRSARVPLRUPRAVRPKKAAWAPPVIDVRLALPAGGPV
AEALASAAVAGCFOVVGHVDSLSVALRHVAATARAATPEPEVEVNGDENG
EHWSPGDCQGWAGNWLQSGASHFRNTADALPIQLEQTAKIMDVLRGAVATQS
IAGAUTNGSLCIRKRRRQDDRGSGSGPIRHDDILRLMLRSSRCSRALALHLCPGA
SAFHLPFRSGFRPMGAVVTVGDQACNGLYKSVAGPKAVSNDLDRNGDGT
GVASAEFLYCFPSAGTAKGASEVLSADAGKIPLNLQFWAAACVLGYHFLSLHS
TWLL"
complement(join(31081. 31242,31391. 31579,31787. 31815,
31903. 31993,32132. 32306,32410. 32549,32638. 32770,
33508. 33574,34086. 34243,34332. 34405,34496. 34609,
35132. 35437)
/gene="OSUNBa0014K08.6"
complement(join(31081. 31242,31391. 31579,31787. 31815,
31903. 31993,32132. 32306,32410. 32549,32638. 32770,
33508. 33574,34086. 34243,34332. 34405,34496. 34609,
35132. 35437)
/gene="OSUNBa0014K08.6"
/note="contains ESTs C72949 (E2532), C72949 (E2532)"
/codon start=1
/product="putative casein kinase-like protein"
/protein id="BAC05582.1"
/db_xref="GI:21902033"
/translation="MPVLSAARRAREAOENPAAAAAAPPARRRRRAARKEPEV
AVEAAPAEEGREEIEVADLGRGGEGKMEGPDGSGARSADKQAVEDEGNTVPDPT
VQVSGSPKYVDVKLKGFGQVYVGRMLANGPGASEVALKFEHRTSKGCHGPPYE
WQYNAIGGHPVRHYKRGQEIYIVTTFDIYERNCFLFALSKLINLVIEKS
TKWRDASTGLHYDQRPVDFREYVYVNLKPEEPYAKICISLDTVGPNDIRP
INVDGAQLVHIDKNGEDGFIISCTSCSNLWALIMDAGTGFQSVTELPFLHKEWI
MQWRNRYITALAGANNSSLVMSKGTMYTQSVKVVLEDFLYPSEGHORWDNGY
RITATAWDOAAFLVSVPRKPTDETQETLTSAPPSQHVKSGLHVAISRLLA
LVSAQSAAYVMLASSEDLAIPDYRAVVKIFNGYS"
join(39834. 39907,40580. 40722,41299. 41408)
/gene="OSUNBa0014K08.7"
join(39834. 39907,40580. 40722,41299. 41408)
/gene="OSUNBa0014K08.7"
/note="hypothetical protein"
/codon start=1
/protein id="BAC05583.1"
/db_xref="GI:21902034"
/translation="MPYVVTGSGNSEAIVIRSSLSIRQKEEYPKGYKSELANPPM
AVGPSLDAKSCTFKHICRDYSGERTLCVSVIENRRFAPVSLDVQNKREVVVLCGLDI
SIFPPC"

Query Match 18.3%; Score 32.4; DB 8; Length 177263;
Best Local Similarity 53.1%; Pred. No. 71;
Matches 69; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 9 GTGACGAGCGCAGCTTCTGTGTGCTAACCGGAGCGGAGCGAACTCTTATCGAA 68
DB 148416 GTGACGAGCGAGCGTACGTAATGTGGTGGAAATCTCGCCCGGTGCAAAATCTTGTGGCGCA 148475
QY 69 CAGGACGGCCCTCCATATCAGCGCGCGTATCTCATGCGGTGACCGGACACGAGCGCG 128
DB 148476 AAGGCGCTCGGCTTGTTCGGGATTCGCTGTCACGCTGCGCCCGCGCGCGCC 148535
QY 129 CGGTCCCGCT 138
DB 148536 CCGGCGCTCT 148545
RESULT 11
AC134482/c 222256 bp DNA linear HTG 15-NOV-2002
LOCUS AC134482

DEFINITION	Rattus norvegicus clone CH230-179K4, *** SEQUENCING IN PROGRESS
ACCESSION	***, 2 unordered pieces.
VERSION	AC134482
KEYWORDS	HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 222256)
AUTHORS	Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Kocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheewa, L., Loulseghe, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackelmech, O., Okwou, G., Olarunpasegbon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L. L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.
REFERENCE	2 (bases 1 to 222256)
AUTHORS	Rat Genome Sequencing Consortium.
TITLE	Submitted (27-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
JOURNAL	3 (bases 1 to 222256)
REFERENCE	Rat Genome Sequencing Consortium.
AUTHORS	Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One
JOURNAL	
COMMENT	Baylor Plaza, Houston, TX 77030, USA On Nov 15, 2002 this sequence version replaced gi:23334632. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
	Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu Project Information Center project name: KCOV Center clone name: CH230-179K4 Summary Statistics Assembly program: Phrap; version 0.990329 Consensus quality: 190050 bases at least Q40 Consensus quality: 194119 bases at least Q30 Consensus quality: 196964 bases at least Q20 Estimated insert size: 204504; sum-of-contigs estimation Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
	* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html). * NOTE: This is a 'working draft' sequence. It currently * consists of 2 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved.
	* 1 107899: contig of 107899 bp in length * 107900 107999: gap of unknown length * 108000 222256: contig of 114257 bp in length.
FEATURES	Location/Qualifiers 1..222256 /organism="Rattus norvegicus" /mol_type="genomic DNA" /db_xref="taxon:10116" /clones="CH230-179K4" complement(113301..114059) /note="clone boundary clone_end:Sp6 site: end_sequence:BH303478"
misc_feature	
ORIGIN	Query Match 18.3%; Score 32.4; DB 2; Length 222256; Best Local Similarity 53.1%; Pred. No. 71; Matches 69; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY	44 AGCCGGACGCAATCTCTTATCGGAACAGGACGCGCTCCATATCAGCGCGCGTTATCTC 103
Db	2737 AACTAGACCCCGCTCCTCATCAGCTCAGTACATATCTCTCTGGGCTGGACCCCTTCTATC 2678
QY	104 ATGCGGTGACCGACACGAGCGCGCTTCCTTCGCGCTTATCGGCTTATAAATACAGCCGCGC 163
Db	2677 GTTAGCCAGAGTACAGAACAGGCTCAGTCCCTTGTTCACACTTCTTAAGGAAGGTAC 2618
QY	164 AACGATCTGG 173
Db	2617 AGAGTGCTGG 2608

REFERENCE AUTHORS

QY 44 AGCCGACGCACTCTTATCGAACAGAGCGGCTCCATATACGACGCGGTATATCTC 103
Db 105760 AACTAGACCCGCTCTCATCAGCTCAGTACATCTCTCTGGCGCTGACCCCTTCTATC 105819
QY 104 ATGCGCTGACCGGACACAGAGCGGCTCGGCTTATCGGCTTATCGGCTTATAAATACAGCCGCG 163
Db 105820 GTTAGCCAGAGTGAGACAGGCTCAGGTCCCTTGTTCACACTTCTTAGGAGAGTCCAC 105879
QY 164 AACGATCTGG 173
Db 105880 AGAGTGCTGG 105889

RESULT 13
AP005031
LOCUS Streptomyces avermitilis MA-4680 linear BCT 10-MAY-2003
DEFINITION Streptomyces avermitilis genomic DNA, complete genome, section 11/30.
ACCESSION AP005031 BA000030
VERSION AP005031.1 GI:29606108
KEYWORDS
SOURCE Streptomyces avermitilis MA-4680
ORGANISM Streptomyces avermitilis MA-4680
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

1
REFERENCE
AUTHORS Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., Kikuchi, H., Shiba, T., Sakaki, Y., and Hattori, M.
TITLE Genome sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary metabolites
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (21), 12215-12220 (2001)
MEDLINE 21477403
PUBMED 11572948

2
REFERENCE
AUTHORS Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M., and Omura, S.
TITLE Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis
JOURNAL Nat. Biotechnol. 21 (5), 526-531 (2003)
MEDLINE 22608306
PUBMED 12692562

3 (bases 1 to 299550)
REFERENCE
AUTHORS Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., Kushida, N., Director-General of Biotechnology Center, Shiba, T., Sakaki, Y., and Hattori, M.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2002) Director-General of Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; 2-49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan (E-mail: bio@nite.go.jp, URL: http://www.bio.nite.go.jp/, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424)
COMMENT This work was done in collaboration with Haruo Ikeda(*1), Jun Ishikawa(*2), Akiharu Hanamoto(*3), Chigusa Takahashi(*3), Mayumi Shinose(*3), Hiroshi Horikawa(*4), Hidekazu Nakazawa(*4), Tomomi Osonoe(*4), Norihiro Kushida(*4), Hisashi Kikuchi(*4), Tadayoshi Shiba(*5), Yoshiyuki Sakaki(*6,*7), Masahira Hattori(*1,*7) and Satoshi Omura(*1,*3).
Final finishing process and all annotation were done by H. Ikeda and J. Ishikawa.
*1 Kitasato Institute for Life Sciences, Kitasato University
*2 National Institute of Infectious Diseases
*3 The Kitasato Institute
*4 National Institute of Technology and Evaluation
*5 School of Science, Kitasato University
*6 Institute of Medical Science, University of Tokyo
*7 RIKEN, Genomic Sciences Center
Following url is also available.
http://avermitilis.is.kitasato-u.ac.jp.
Location/Qualifiers

source

1. 299550
/organism="Streptomyces avermitilis MA-4680"
/mol_type="genomic DNA"
/strain="MA-4680"
/db_xref="taxon:227882"
/note="This strain is also named as strain: ATCC 31267, NCIMB 12804 or NRRL 8165.-synonym: Streptomyces avermectinius"
complement(42..2936)
/gene="nrd"
complement(42..2936)
/gene="nrd"
/note="SAV2461"
/codon_start=1
/transl_table=11
/product="putative vitamin B12-dependent ribonucleotide reductase"
/protein_id="BAC70172.1"
/db_xref="GI:29606109"
/translation="MTETASGPARGSPAKCTKAKGLIERIHTTPGVHPYDEVWEARR DVVNNWRDSVNFQRGVPPDFWSVNAVNIUTSKYFRGAVGTPQREVSLRQLIDRI VKTRYKAGEDYKYFASPADAEIFEHELAYALLHQIFSNFVFNVTGTPQPVQSACF ILAVDDSMESILDWYKEGMIFKGSGAGLNLRISSKELLSSGSGNASGVPSPMRGA DASAGTIKSGGATRAAKMVLVDVHPDIEDFIETVKVEBEKIRALRDAGDMDLGGD DITSVOYONANNSVRVNDTFMKAVEGKGFLTSRMTGEVIEVDKSLFKMAEAAW GRADPGIOYDTHNHMTCPESGRINGSNPCEVHLDNTSCNLASLNMKFLKDDGK ACADPGVERFAKVELVITAMDISIICPADPPTQKICENTRAFQRLGIGYANILGALJA TGHAYSDSGKRALAGATISLMTGTSYKRSAAELAAVGYDYGIANKAGPHORVMKQHS ALGVAVRVDDLDTPIWAAATEAWQDVLHGEKNGFRNAQSVIAPTGTIIGLMSCDTT GPEPDIALVKFKLVLGSGMQIVNGTVQALRGLYQEQEIEAIVAHIAENGVIDVY FLKHEHYEVDFDCAMGERSISAGHVRMMAAIQPMWISGALNTVNLPTATVDEEYV FLKWMGSKALAIYRDNCKVGOPLSAKKETEKAEVTAKEATIREAVEKVEVRYVR KRIPKRGPIITTSFTVGGAGYMTANSYPDDGLGVEFLKMSKOGSTLAGMWDASIAV SVGLYGVPLETIVTSKFTNMRFPAGMTDDPDMVMAQSIYDIIFRRLALDFLPETRS ALGHTSAELVEMQLGIQADAPLFCSCGTMRAGSGYICEGCGSTSGCS"
complement(3086..3649)
/note="SAV2462"
/codon_start=1
/transl_table=11
/product="putative regulatory protein"
/protein_id="BAC70173.1"
/db_xref="GI:29606110"
/translation="MHCPFCRHPDSRVVDSKRTDDGTSIRRRQCPDCSRRTTVEIC SLMVVKESGVTETPEFSTKINGVKACQGVTEDEALQIGQVVEEAVRATGSAELTT HDVGLAILGPERELDLVAYLRPASVYRPAFSLDEFEAAIAELREQRERPAVDDDEH DAGARQGTDRGSGGTVEVPVATVAD"
4168..4947
/gene="lexA"
4168..4947
/gene="lexA"
/note="SAV2463"
/codon_start=1
/transl_table=11
/product="putative SOS regulatory protein LexA"
/protein_id="BAC70174.1"
/db_xref="GI:29606111"
/translation="MTTASATITAGDSQGRLEPVMANNEATSHGPKRSLGRPP GIRADSSGLDRQRVRVLEIRDSVQRGYPSPMRKEIGQAVGLSTSSVAHQLMALERK GFRLDPHPRAYEVRGSDSSQVPTDTAGKPAASVPLVGLRIAGAGPTLAEESEVD VFPLRQLVGDGELFVLKVGDSMIEAAICDGMWTVTRQFPVKAENGDIVAAMLDEGAT VKRFRDGHVWLLPHNSAYQIPGDEATILGKVVALRVV"
complement(5110..7143)
/gene="ding"
complement(5110..7143)
/gene="ding"
/note="SAV2464"
/codon_start=1
/transl_table=11
/product="putative ATP-dependent helicase"
/protein_id="BAC70175.1"
/db_xref="GI:29606112"
/translation="MTKPSLPFLHAAVATVAGGTERPQGVMTAEAVAEIDSSHLV

OACTGCTCKSLGVLPAALHGERVIVATATLALQOLVERDLPRIVDALHPLRRRPEF
AMUKGRNYSCLRWQHSVSPQDEEGFLDQFEAAAPTSLKGLDLRLRDSWDETETGD
RDLTLPVSDRAHQAQVSGVEEGFLDQFEAAAPTSLKGLDLRLRDSWDETETGD
IDAI EGAPVLPQHEVLIVDEAHELVSVTGATCELTTPGVNRAVRRAAKLVNEKAA
QLOTAAGPFLRLMELAPGLEBI PEDIGVALMALRDAARTVISAIGADTKSVQDED
AVRKOALASVESVHDVAERTTNGSEWDVWYVHERDRFGASLRVAPMSVGLLREKLP
DRSVLTATSLKGLPENGVCASIGLAPGEGTQDDLPQWKGVDVFPDFYKPGILYV
AKHLSRPARDQDMDDELITELIQAGGRTGLFSMRQAQALAEELSRIPPEFPIV
LQOEETLGEILKNFADVPETCLFTLSLMQGVDPVPGPSCOLVMDKIPFRPDPLMS
ARQKAVEDAGNGFMVAATHAALLMAQAGRLVRASGDRGVAVLDLQRLATARYGSY
LKASLPDFTYTRDNRVRSRLTAIDEAAKTEVTEVAEVEEAAKESEGA"
complement (7376..8509)
/note="SAV2465"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC70176.1"
/db_xref="GI:29606113"
/translation="MPPTDASTDAEPATSDAPARIASARSADSEDTLELRPLDELV
ALFAAEAGEEBAAGGGGAVRGEETGAGGEDAVGGAEPEAGRKDAGPEAGTATA
APSPARAGATAASDSCGSDAVDPAAPFAAASTVPSASRPSVTSAPSVTAT
APAPAPSLGDTKALAGQAVSVTRAADDLLGRVEDWGTPTVPAQTFLQAPVRRDL
PLLRWMDPAVAAPFKLAGDSVTEQHILRAOLGGDGRSVPCIGVLEGTMPMSWEIYR
ADLSLARHYPARPHGTGTHLJGGVADRGRLGSLTLLRAVADLVLDPRPCARVAE
PDLRNTSSVAFICAGFRSAEVDLFDKRAALMVRDRPLRDL"
complement (8549..10609)
/note="SAV2466"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC70177.1"
/db_xref="GI:29606114"

CDS

/translations="MNATPTSDGSRSTHERAGHDTQLTPTVQGPQAVFASQAAPAE
TOAPLAPATOVVPEVTPVVOASVSPVPPVGRASERLCTTADPLDHPAOTAA
QAAANLLCVRVREGLVAPDPTGLRI PLLATGTALLVPVHWSPTGWRGLPLVLA
DAPQTAPPADVPVALLAREAGSRAPRAPSAVDHCAHDSADLVGRVADSVRR
AVFIDRRRAPADPEPFLAEEQVALLHGLPHTPKSRGLSEARLYSPELGRSFA
LHLVAVPPLATNSAWTERGREVPAPQOLTRAGLDGLPMDPGFAALPLHPQIREIR
HREPAALDLAGIIRDLGPHGPAPHTPSVTRVHTCAAPAMKLISGLRITNSRREL
RKLHRTGVEVHRLSLGSLGEOMQAAHPGPDIVRDPAPLAVDDOHNAPGLDVMIRHN
PPAPTDVSCVAGLSPWPPOPEQEPVPMRSRLAEIILRLAGTRGRGAVATWF
LRYLQVVRVRLMDSEAGIALSAHQNTLLLDPEGWTFPGYRNDNQGYFTFSRRA
ELRPLPGTGEHRTFVSEVDETFERFAYLGINNVGLIGAFSGHLAEQLELLAAFR
RFLADVATGPARLRTPLRLLSPVLRCKANLLTRHLGLDELVGFDVTSVVYTIAN
PLHS"
complement (10775..12229)
/note="SAV2467"
/codon_start=1
/transl_table=11
/product="putative class III aminotransferase"
/protein_id="BAC70178.1"
/db_xref="GI:29606115"
/translation="MATVKSAYASLKRBAEAGHAHEGILRQSARSAEARTYARALP
IVPVRAGLTIBAGDRYLDLCUSGAGTIALGHNHPVLEAIRKVIDSGAPLHVLDA
TPVKATFELTFLPAGLADRARVQFCGPAGTDAVEAAFKLVRTATGRSGLLAFTGA
YHGMTAGLEASGATGRVARLPYQDYCPFGICIGERGAEIGARWESLLDDVKSQ
VPLPAGMILEPVOEGGVIPAPDAWLRMRATADRSIPLIADVOTGVGRTGRFVAV
EHSVVDPVVMVLSKISGLSPLAVVVRDLDVMPQAGAHGTFRGNQLAMAGATLA
YVREHLAERAQTLGARMLDQLSLATEPFCVGEVRGRLMIGVELVRAEDREAVSP
GSGDFAADTAGAGPAPAPAPLAAAVQRECLRRGLLIVELGGRHASVRLPLPLTISDE
QATVADRLDLAGAVEAVARGAERSRGRFQSERAG"
12761..13993
/note="SAV2468"
/codon_start=1
/transl_table=11
/product="putative secreted protein"
/protein_id="BAC70179.1"
/db_xref="GI:29606116"
/translations="MRPIRLPFAARRGRARRRTSPVLAGVGLATALTATACDSGD
DTKASSESPSASAGDDKIKIPDDIRDLKIERGIDIDKWKNGAWKWRDWDLEARN
18.2%; Score 32.2; DB 1; Length 299550;
Best Local Similarity 61.2%; Pred. No. 80;

CDS

CDS

10775..12229
/note="SAV2467"
/codon_start=1
/transl_table=11
/product="putative class III aminotransferase"
/protein_id="BAC70178.1"
/db_xref="GI:29606115"
/translation="MATVKSAYASLKRBAEAGHAHEGILRQSARSAEARTYARALP
IVPVRAGLTIBAGDRYLDLCUSGAGTIALGHNHPVLEAIRKVIDSGAPLHVLDA
TPVKATFELTFLPAGLADRARVQFCGPAGTDAVEAAFKLVRTATGRSGLLAFTGA
YHGMTAGLEASGATGRVARLPYQDYCPFGICIGERGAEIGARWESLLDDVKSQ
VPLPAGMILEPVOEGGVIPAPDAWLRMRATADRSIPLIADVOTGVGRTGRFVAV
EHSVVDPVVMVLSKISGLSPLAVVVRDLDVMPQAGAHGTFRGNQLAMAGATLA
YVREHLAERAQTLGARMLDQLSLATEPFCVGEVRGRLMIGVELVRAEDREAVSP
GSGDFAADTAGAGPAPAPAPLAAAVQRECLRRGLLIVELGGRHASVRLPLPLTISDE
QATVADRLDLAGAVEAVARGAERSRGRFQSERAG"
12761..13993
/note="SAV2468"
/codon_start=1
/transl_table=11
/product="putative secreted protein"
/protein_id="BAC70179.1"
/db_xref="GI:29606116"
/translations="MRPIRLPFAARRGRARRRTSPVLAGVGLATALTATACDSGD
DTKASSESPSASAGDDKIKIPDDIRDLKIERGIDIDKWKNGAWKWRDWDLEARN
18.2%; Score 32.2; DB 1; Length 299550;
Best Local Similarity 61.2%; Pred. No. 80;

CDS

Query Match
Best Local Similarity 61.2%; Pred. No. 80;

Matches	52;	Conservative	0;	Mismatches	33;	Indels	0;	Gaps	0;
Qy	58	CCTTATCGGAACAGGCGCCCTCCATATCAGCGCGCGTTATCTCATGCGCGTGACCGG	117	129986	CCTCCACCGAACCGGACGCGTCCCTTGTATGATGAGTTGAGTTCTCTGCACGACCGG	129045			
Db	129986	CCTCCACCGAACCGGACGCGTCCCTTGTATGATGAGTTGAGTTCTCTGCACGACCGG	129045						
Qy	118	ACACGAGGCGCGCCGCTTCCTTCCTTATC	142	129046	CCTTACGCGCGCTCGTCCAGGTTCTC	129070			
Db	129046	CCTTACGCGCGCTCGTCCAGGTTCTC	129070						
RESULT 14									
BD221352/c	BD221352	Human gene and gene expression product V.	921 bp	DNA	linear	PAT 17-JUL-2003			
LOCUS	BD221352	Human gene and gene expression product V.							
DEFINITION	BD221352	Human gene and gene expression product V.							
ACCESSION	BD221352	Human gene and gene expression product V.							
VERSION	BD221352.1	GI:33031122							
KEYWORDS	JP 2002534055-A/2465.								
SOURCE	JP 2002534055-A/2465.	Homo sapiens (human)							
ORGANISM	Homo sapiens								
REFERENCE	1 (bases 1 to 921)	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
AUTHORS	Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.S., Reinhard, C., Giese, K., Randazzo, F., Kennedy, G.C., Pot, D., Kaseam, A., Lanson, G., Drmanac, R., Crkvenjakov, R., Dickson, M., Drmanac, S., Labat, I., Leshkowitz, D., Kita, D., Garcia, V., Jones, L.W. and Crain, B.S.								
TITLE	Human gene and gene expression product V								
JOURNAL	Patent: JP 2002534055-A 2465 15-OCT-2002;								
COMMENT	CHIRON CORP, HYSEQ INC								
	OS Homo sapiens (human)								
	PN JP 2002534055-A/2465								
	PD 15-OCT-2002								
	PP 13-MAY-1999 JP 2000548466								
	PR 14-MAY-1998 US 60/085426, 15-MAY-1998 US 60/085537 PR								
	15-MAY-1998 US 60/085696, 21-OCT-1998 US 60/105234 PR								
	27-OCT-1998 US 60/105877								
	PI LOUIS T WILLIAMS, JAIME ESCOBEDO, MICHAEL A INNIS, PABLO PI DOMINGUEZ GARCIA.								
	PI JULIE SUDDUTH KLINGER, CHRISTOPH REINHARD, KLAUSE GIESE, FILIPPO RANDAZZO,								
	PI GIULIA C KENNEDY, DAVID POT, ALTAF KASSAM, GEORGE LAWSON, RADOJE DRMANAC,								
	PI RADOMIR CRKVENJAKOV, MARK DICKSON, SNEZANA DRMANAC, IVAN LABAT,								
	PI DENA LESHKOWITZ, DAVID KITA, VERONICA GARCIA, LEE WILLIAM JONES,								
	PI BIRGIT STACHE CRAIN								
	PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21 PC								
	, C12N5/00, C12Q1/68,								
	PC C12N15/00, C12N5/00								
	CC n = A, T, C or G								
	FF Key	Location/Qualifiers							
	FT misc feature (1)..(921).	Location/Qualifiers							
	source	1..921							
		/organism="Homo sapiens"							
		/mol_type="genomic DNA"							
		/db_xref="taxon:9606"							
ORIGIN									
Query Match	18.1%;	Score 32;	DB 6;	Length 921;					
Best Local Similarity	51.5%;	Pred. No. 1.1e+02;							
Matches	68;	Conservative	0;	Mismatches	64;	Indels	0;	Gaps	0;
Qy	16	GAGCGCAGCTTCTGTGTGTTGCTTAACCGCAGCGGACGCACTCTTATCGGAACAGACG	75	157	GGCAGCTGCGCNCAGGTGGCTGTCTCCACGCGCGTCTCCCGCGTCCCGCGGT	98			
Db	157	GGCAGCTGCGCNCAGGTGGCTGTCTCCACGCGCGTCTCCCGCGTCCCGCGGT	98						
Qy	76	CGCTCCATATCAGCGCGCGTTCATCTCATGCGCGTACCGGACGACGACGCGCGCTCCC	135	97	GGCGTCNAANCACCTTCCCTTCCTTCCTCCGGTCCCGGTCAGGAGCGACTGTCTCT	38			
Db	97	GGCGTCNAANCACCTTCCCTTCCTTCCTCCGGTCCCGGTCAGGAGCGACTGTCTCT	38						
Qy	136	GCTTATCGCGCC	147						

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 01:20:43 ; Search time 193.559 Seconds
(without alignments)

5413.313 Million cell updates/sec

Title: US-09-896-888a-1_COPY_351_527

Perfect score: 177

Sequence: 1 gctttatctgacagacgc.....gcccgcaacgatcttgtaaa 177

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 segs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	177	100.0	560	ADQ48575	Adq48575 OpIE2 pro
2	177	100.0	564	AAV62487	AAV62487 O. pseudo
3	177	100.0	2773	AAAL61306	AAAL61306 p2Op2F e
4	175.4	99.1	5038	ADQ48539	Adq48539 Viral vec
5	41	23.2	141	ADQ48576	Adq48576 Viral vec
6	32	18.1	921	AAA02474	AAA02474 Human col
7	30.8	17.4	1268	ABL18801	ABL18801 Bacterial
8	30.8	17.4	1342	ABL27503	ABL27503 Drosophil
9	30.8	17.4	3448	ABL27502	ABL27502 Drosophil
10	30.8	17.4	3600	ABL18800	ABL18800 Drosophil
11	30.8	17.4	29169	AA064735	AA064735 Human car
12	30.6	17.3	1404	AD147065	AD147065 Bacterial
13	30	16.9	657	AD145632	AD145632 Plant iso
14	30	16.9	1710	ABD14088	ABD14088 Pseudomon
15	30	16.9	3107	AD889385	AD889385 Oligonuc
16	30	16.9	3501	AD889483	AD889483 Oligonuc
17	30	16.9	6107	ABK31430	ABK31430 Signal tr
18	30	16.9	6107	ABL70389	ABL70389 Chemical
19	30	16.9	6107	AA61341	AA61341 Human gen
20	29.8	16.8	1370	ADF83430	ADF83430 Bread whe

21	29.8	16.8	1789	4	ABL19245	Ab119245 Drosophil
22	29.8	16.8	3915	4	ABL19244	Ab119244 Drosophil
23	29.6	16.7	300	3	AAA01209	AAA01209 Human col
24	29.2	16.5	1593	13	ADT45695	Adt45695 Bacterial
25	28.6	16.2	585	11	ABD13766	Abd13766 Pseudomon
26	28.6	16.2	823	12	ACH89781	Ach89781 Human gen
27	28.6	16.2	915	5	AA570780	Aas70780 DNA encod
28	28.6	16.2	1425	4	AA551490	Aas51490 Pseudomon
29	28.6	16.2	1425	8	ACA19456	AcA19456 Prokaryot
30	28.6	16.2	1434	13	ADS56860	Ads56860 Bacterial
31	28.6	16.2	1443	11	ABD13493	Abd13493 Pseudomon
32	28.6	16.2	1551	11	ABD13719	Abd13719 Pseudomon
33	28.6	16.2	2085	11	ABD13633	Abd13633 Pseudomon
34	28.6	16.2	2091	2	AA06989	Aax06989 Human neu
35	28.6	16.2	2091	6	ABK90730	Abk90730 cDNA enco
36	28.6	16.2	2502	2	AA06988	Aax06988 Human neu
37	28.6	16.2	2502	6	ABK90731	Abk90731 cDNA enco
38	28.4	16.0	876	13	ADT43165	Adt43165 Bacterial
39	28.4	16.0	1320	8	ACA45469	Aca45469 Prokaryot
40	28.4	16.0	29040	11	ADL27170	Adl27170 Human gen
41	28.4	16.0	29322	9	ADA03092	Ada03092 Human LFN
42	28.4	16.0	29322	9	ADA66376	Ada66376 Human LFN
43	28.4	16.0	29322	10	ADB72830	Adb72830 Human LFN
44	28.4	16.0	75839	11	ACN43958	Acn43958 Human gen
45	28.2	15.9	283	3	AAC23731	Aac23731 Human sec

ALIGNMENTS

RESULT 1
ADQ48575
ID ADQ48575 standard; DNA; 560 BP.
XX
AC ADQ48575;
XX
DT 09-SEP-2004 (first entry)
XX
DE OpIE2 promoter DNA sequence.
XX
KW viral vector; recombination site; recombinant virus;
KW replication-defective particle generation; gene expression inhibition;
KW gene therapy vector; ds; OpIE2 promoter.
XX
OS Unidentified.
XX
PN WO2004009768-A2.
XX
PD 29-JAN-2004.
XX
PF 18-JUL-2003; 2003WO-US022437.
XX
PR 18-JUL-2002; 2002US-0396335P.
PR 26-JUL-2002; 2002US-0398617P.
PR 19-NOV-2002; 2002US-0427231P.
PR 24-MAR-2003; 2003US-0456496P.
PR 03-JUN-2003; 2003US-0474940P.
XX
PA (INVI-) INVITROGEN CORP.
PA (BENN/) BENNETT R P.
PA (WELC/) WELCH P J.
PA (HARW/) HARWOOD S.
PA (MADD/) MADDEN K.
PA (FRIM/) FRIMPONG K.
PA (FRAN/) FRANKE K E.
XX
PI Bennett RP, Welch PJ, Harwood S, Madden K, Frimpong K, Franke KB;
XX WPI; 2004-132944/13.
XX
PT Novel nucleic acid molecule comprising all or a portion of one or more
PT viral genome and further comprising two or more recombination sites that
PT do not substantially recombine with each other, useful as gene therapy.

XX Disclosure; Fig 16; 555pp; English.

PS The invention comprises a nucleic acid molecule consisting of all or a

XX portion of one or more viral/baculoviral genomes and further containing

CC at least two recombination sites that do not substantially recombine with

CC each other. One or more of the recombination sites is capable of

CC undergoing recombination with a compatible recombination site in the

CC presence of one or more proteins active in lambda recombination. The

CC nucleic acid of the invention replicates in prokaryotic and eukaryotic

CC cells. The nucleic acid of the invention is useful for constructing a

CC recombinant virus, generating replication-defective particles,

CC preventing/inhibiting the expression of one or more genes in an organism,

CC and are useful as gene therapy vectors. The nucleic acid of the invention

CC is also useful for producing and expressing fusion polypeptides. The

CC present DNA sequence represents the OplE2 promoter that was used in the

CC exemplification of the invention.

XX Sequence 560 BP; 144 A; 156 C; 116 G; 144 T; 0 U; 0 Other;

SQ

Query Match 100.0%; Score 177; DB 12; Length 560;

Best Local Similarity 100.0%; Pred. No. 2.6e-48;

Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCTTATCGTGACAGAGCCAGCTTCCTGTTGCTTAACCGCAGCGCGGCAACTCCT 60

DB 355 GTCTTATCGTGACAGAGCCAGCTTCCTGTTGCTTAACCGCAGCGCGGCAACTCCT 414

QY 61 TATCGGAACAGGACGCGCTCCATATCAGCGCGGCTTATCTCATGCGGTGACCGGACA 120

DB 415 TATCGGAACAGGACGCGCTCCATATCAGCGCGGCTTATCTCATGCGGTGACCGGACA 474

QY 121 CGAGGCGCGCGTCCCGCTTATCGCGCTTATAATACAGCGCGGCAACTCTGGTAAA 177

DB 475 CGAGGCGCGCGTCCCGCTTATCGCGCTTATAATACAGCGCGGCAACTCTGGTAAA 531

RESULT 2

AAV62487

ID AAV62487 standard; DNA; 564 BP.

AC AAV62487;

XX 17-OCT-2003 (revised)

DT 19-JAN-1999 (first entry)

XX

DE O. pseudotsugata multicapsid nucleopolyhedrosis virus ie2 promoter.

XX

XX Orgyia pseudotsugata; multicapsid; nucleopolyhedrosis virus; OpNPV;

KW Op ie2; promoter; shuttle vector; transformation; melanotransferrin;

KW immediate early baculovirus promoter; prokaryotic; transcription;

KW bleomycin/bleomycin-type antibiotic; insect cell; transposon;

XX ion transport peptide hormone; ss.

XX

OS Orgyia pseudotsugata; polyhedrosis virus.

XX

XX WO9844141-A2.

PN

XX

PD 08-OCT-1998.

XX

PF 26-MAR-1998; 98WO-CA000282.

XX

PR 27-MAR-1997; 97US-0049946P.

PR 28-JAN-1998; 98CA-02221819.

XX

XX (UYBR-) UNIV BRITISH COLUMBIA.

PA

XX Grigliatti TA, Theilmann DA, Pfeifer TA, Hegedus DD;

PI

XX WPI; 1998-557129/47.

XX

DR Expression vectors for transforming insect cells from disparate lines -

XX useful to express heterologous DNA, e.g. to allow study of gene

PT

expression and produce commercially important proteins.

Claim 10; Page 82; 121pp; English.

This represents a Orgyia pseudotsugata multicapsid nucleopolyhedrosis virus (OpNPV) immediate early 2 (Op ie2) promoter sequence. The invention provides a new shuttle vector for transforming insect cells that comprises: (i) prokaryotic origin of replication; (ii) insect promoter having homology to, and capable of functioning as, an immediate early baculovirus promoter; (iii) prokaryotic promoter sequence, and (iv) selectable marker capable of conferring resistance to a bleomycin/bleomycin-type antibiotic under transcriptional control of (ii) and (iii), in insect and prokaryotic cells respectively. The vectors can be used to stably transform (especially insect) cells with heterologous DNA, useful to allow study of gene expression and direct expression of heterologous gene products e.g. commercially important proteins. They are especially useful to allow expression of the heterologous melanotransferrin, ion transport peptide hormones or biologically active derivatives in insect cells. They enable screening of lines for optimum lines from disparate species, allowing screening of lines for optimum post-translational modification of particular proteins. Shuttle vectors further comprising DNA transposable elements defining a transposon can be used to optimise heterologous protein expression and facilitate selection of desired transformants. (Updated on 17-OCT-2003 to standardise OS field)

XX

SQ Sequence 564 BP; 147 A; 160 C; 116 G; 141 T; 0 U; 0 Other;

Query Match 100.0%; Score 177; DB 2; Length 564;

Best Local Similarity 100.0%; Pred. No. 2.6e-48;

Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCTTATCGTGACAGAGCCAGCTTCCTGTTGCTTAACCGCAGCGCGGCAACTCCT 60

DB 351 GTCTTATCGTGACAGAGCCAGCTTCCTGTTGCTTAACCGCAGCGCGGCAACTCCT 410

QY 61 TATCGGAACAGGACGCGCTCCATATCAGCGCGGCTTATCTCATGCGGTGACCGGACA 120

DB 411 TATCGGAACAGGACGCGCTCCATATCAGCGCGGCTTATCTCATGCGGTGACCGGACA 470

QY 121 CGAGGCGCGCGTCCCGCTTATCGCGCTTATAATACAGCGCGGCAACTCTGGTAAA 177

DB 471 CGAGGCGCGCGTCCCGCTTATCGCGCTTATAATACAGCGCGGCAACTCTGGTAAA 527

RESULT 3

AAV61306

ID AAV61306 standard; DNA; 2773 BP.

AC AAV61306;

XX 22-SEP-2003 (first entry)

DT

XX p2Zop2F expression vector for insect cells.

DE

XX Multimeric protein; interleukin 5; IL5; TNFalpha; inflammatory disease;

KW tumour necrosis factor alpha; gene therapy; arthritis; gene; ds.

KW

XX Unidentified.

OS

XX WO2003042244-A2.

PN

XX

PD 22-MAY-2003.

XX

PF 15-NOV-2002; 2002WO-DK000764.

XX

PR 16-NOV-2001; 2001DK-00001702.

PR 16-NOV-2001; 2001US-0331575P.

XX

XX (PHAR-) PHARMEXA AS.

PA (KLYS/) KLYSNER S.

PA (NIEL/) NIELSEN F S.

PA (BRAT/) BRATT T.

PR 03-JUN-2003; 2003US-0474940P.
XX (INVI-) INVITROGEN CORP.
PA (BENN/) BENNETT R P.
PA (WELC/) WELCH P J.
PA (HARW/) HARWOOD S.
PA (MADD/) MADDEN K.
PA (FRIM/) FRIMPONG K.
PA (FRAN/) FRANK K E.
XX
PI Bennett RP, Welch PV, Harwood S, Madden K, Frimpong K, Franke KE;
XX WPI; 2004-132944/13.
XX
XX Novel nucleic acid molecule comprising all or a portion of one or more
PT viral genome and further comprising two or more recombination sites that
PT do not substantially recombine with each other, useful as gene therapy.
XX
XX Disclosure; Fig 17; 555pp; English.
XX
XX The invention comprises a nucleic acid molecule consisting of all or a
CC portion of one or more viral/baculoviral genomes and further containing
CC at least two recombination sites that do not substantially recombine with
CC each other. One or more of the recombination sites is capable of
CC undergoing recombination with a compatible recombination site in the
CC presence of one or more proteins active in lambda recombination. The
CC nucleic acid of the invention replicates in prokaryotic and eukaryotic
CC cells. The nucleic acid of the invention is useful for constructing a
CC recombinant virus, generating replication-defective particles,
CC preventing/inhibiting the expression of one or more genes in an organism,
CC and are useful as gene therapy vectors. The nucleic acid of the invention
CC is also useful for producing and expressing fusion polypeptides. The
CC present DNA sequence represents a recombination region of a plasmid that
CC was used in the exemplification of the invention.
XX
XX Sequence 141 BP; 43 A; 36 C; 27 G; 34 T; 0 U; 1 Other;
SQ

Query Match 23.2%; Score 41; DB 12; Length 141;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 137 CTTATCGCGCCTTAATAATACAGCCGCAACGATCTGGTAAA 177
1 CTTATCGCGCCTTAATAATACAGCCGCAACGATCTGGTAAA 41

RESULT 6
AAA02474/c
ID AAA02474 standard; cDNA; 921 BP.
XX
XX AAA02474;
XX
XX 19-MAY-2000 (first entry)
XX Human colon cancer cell line polynucleotide sequence SEQ ID NO:2465.
XX Human; colon cancer; tumour; diagnosis; gene expression product; probe;
XX detection; cancerous state; metastasis; identification; breast cancer;
XX oestrogen receptor-positive breast cancer; therapy;
XX oestrogen receptor-negative breast cancer; lung cancer; ss.
XX Homo sapiens.
XX
XX WO958675-A2.
XX
XX 18-NOV-1999.
XX
XX 13-MAY-1999; 99WO-US010602.
XX
XX 14-MAY-1998; 98US-0085426P.
XX 15-MAY-1998; 98US-0085537P.
XX 15-MAY-1998; 98US-0085696P.
XX 21-OCT-1998; 98US-0105234P.

PR 27-OCT-1998; 98US-0105877P.
XX (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lanson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
XX WPI; 2000-126369/11.
XX Polynucleotide library used to determine cancerous states of mammalian
XX cells.
XX Claim 1; Page 989; 1097pp; English.
XX
XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
CC libraries constructed from human colon cancer cell lines. The present
CC invention also describes a method of detecting differentially expressed
CC genes correlated with a cancerous state of a mammalian cell, comprising
CC detecting at least one differentially expressed gene product in a test
CC sample derived from a cell suspected of being cancerous, where detection
CC of the differentially expressed gene product is correlated with a
CC cancerous state of the cell from which the test sample was derived. The
CC polynucleotides sequences can be used in a method for detecting
CC differentially expressed genes correlated with a cancerous state of a
CC mammalian cell. The polynucleotides can also be used as probes for
CC detecting and mapping related genes. They can be used in diagnosis and
CC prognosis of diseases and disorders (e.g. identification of pre-
CC metastatic or metastatic cancerous states, stages of cancer, or
CC responsiveness of cancer to therapy). This is particularly for breast
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC negative breast cancer, lung cancer, and colon cancer
XX
XX Sequence 921 BP; 239 A; 214 C; 205 G; 223 T; 0 U; 40 Other;
SQ

Query Match 18.1%; Score 32; DB 3; Length 921;
Best Local Similarity 51.5%; Pred. No. 2.1;
Matches 68; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 16 GACCCGAGCTTCCTGTGTCTTAAACCGCAGCCGACGAACTCTTTATCGGAACAGACG 75
DB 157 GGCAGCTGGCNCAGGTGGCTGTCCACGCGGCTCTCCGCGCTGCGCGGT 98

QY 76 GCGCTCCATATCAGCCGCGGCTTATCTCATCGCGGTGACCGACAGCGCGCCCTCCC 135
DB 97 GCGCTCNAANCCACCTTCCCTTCGTCTCTCCGCTCCGCTCCGAGGCGACTGCTCT 38

QY 136 GCTTATCGCGCC 147
DB 37 TAAGATCGGCC 26

RESULT 7
ABL18801
ID ABL18801 standard; DNA; 1268 BP.
XX
XX ABL18801;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 7876.
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX

```
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 7876; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1268 BP; 295 A; 336 C; 360 G; 277 T; 0 U; 0 Other;
SQ
Query Match 17.4%; Score 30.8; DB 4; Length 1268;
Best Local Similarity 55.7%; Pred. No. 5.6;
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 17 ACGCCAGCTTCCTGTTGTTACTACCGCAGCGGACGCAACTCTTATCGGAACGAGCAGC 76
DB 657 ACAACAGCTTCCTGTTGTTGAGGTGCCACAGCGGGTCTTAGTCTTGGCGGCTCCGGTGGC 716
QY 77 GCCTCCATATCAGCGCGGCTTATCTCATCGCGGTGACCGGACAGC 122
DB 717 GTTGGCACATGGCCATACAAATCTCAAGTCCAGAAAGTACAAG 762
RESULT 8
ABL27503
ID ABL27503 standard; DNA; 1342 BP.
XX
AC ABL27503;
XX
XX 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 33982.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX
XX pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX
XX pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 33979; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1342 BP; 324 A; 347 C; 372 G; 299 T; 0 U; 0 Other;
SQ
Query Match 17.4%; Score 30.8; DB 4; Length 1342;
Best Local Similarity 55.7%; Pred. No. 5.7;
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 17 ACGCCAGCTTCCTGTTGTTGTTACTACCGCAGCGGACGCAACTCTTATCGGAACGAGCAGC 76
DB 731 ACAACAGCTTCCTGTTGTTGAGGTGCCACAGCGGGTCTTAGTCTTGGCGGCTCCGGTGGC 790
QY 77 GCCTCCATATCAGCGCGGCTTATCTCATCGCGGTGACCGGACAGC 122
DB 791 GTTGGCACATGGCCATACAAATCTCAAGTCCAGAAAGTACAAG 836
RESULT 9
ABL27502/C
ID ABL27502 standard; DNA; 3448 BP.
XX
AC ABL27502;
XX
XX 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 33979.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX
XX pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 33979; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
```


RESULT 12
ADT47065
ID ADT47065 standard; CDNA; 1404 BP.
XX AC
XX ADT47065;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #21816.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 45503; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 1404 BP; 236 A; 489 C; 454 G; 225 T; 0 U; 0 Other;

Query Match 17.3%; Score 30.6; DB 13; Length 1404;
Best Local Similarity 62.3%; Pred. No. 6.7;
Matches 48; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 99 ATCTGATGCGGTGACCGGACACGAGCGCCGCTTATCGCGCTTATAATACAG 158
DB 464 ACCGCAAGGCGGTGAACGTCGCGGTGTGTGCTGCGCTTTCGCGCTTTCGCGCAAGG 523
QY 159 CCGCAACGATCTGGTA 175
DB 524 ACCGCGACGATCTGGAA 540

RESULT 13
ADI45632/c
ID ADI45632 standard; DNA; 657 BP.
XX AC ADI45632;
XX
DT 22-APR-2004 (first entry)
XX
DE Plant isoprenoid biosynthesis-associated DNA #6.
XX
KW isoprenoid biosynthesis; ss; isopentenyl diphosphate; IPP;
KW dimethylallyl alcohol; DMAPP; short-chain plastid prenyltransferase;
KW gibberellin; carotenoid; abscisic acid; tocopherol; plastoquinone;
KW phyloquinone; mevalonate pathway; phytosterol; brassinosteroid;
KW ubiquinone; monoterpene; sesquiterpene; protein prenylation; chlorophyll;
KW haeme; yield.
XX
OS Unidentified.
XX
PN US2004010815-A1.
XX
PD 15-JAN-2004.
XX
PF 26-SEP-2002; 2002US-00259194.
XX
PR 26-SEP-2001; 2001US-0325277P.
PR 04-APR-2002; 2002US-0370620P.
PR 04-APR-2002; 2002US-0370743P.
XX
PA (LANG/) LANGE B M.
PA (GHAS/) GHASSEMIAN M.
PA (BRIG/) BRIGGS S P.
PA (COOP/) COOPER B.
PA (GLAZ/) GLAZERROOK J.
PA (GOFF/) GOFF S A.
PA (KATA/) KATAGIRI F.
PA (KREP/) KREPS J.
PA (MOUG/) MOUGHAMER T.
PA (PROV/) PROVART N.
PA (RICK/) RICHKE D.
PA (ZHUT/) ZHU T.
XX
PI Lange BM, Ghassseman M, Briggs SP, Cooper B, Glazebrook J;
PI Goff SA, Katagiri F, Kreps J, Moughamer T, Provart N, Ricke D;
PI Zhu T;
XX
XX WPI; 2004-090562/09.
XX
DR New isolated polynucleotides and polypeptides associated with isoprenoid
PT synthesis in plants, useful for producing transgenic plants, for targeted
PT gene disruption, as well as markers or probes.
XX
PS Disclosure; SEQ ID NO 563; 117pp; English.
XX
CC The invention relates to a polynucleotide (or its complement, protein
CC encoding fragment or reverse complement), comprising a nucleotide
CC sequence encoding a polypeptide comprising an amino acid sequence
CC involved in or associated with the biosynthesis of isoprenoids in a rice
CC plant. Also included are an isolated polypeptide involved in or
CC associated with the biosynthesis of isoprenoids in a plant, an expression

CC cassette comprising the polynucleotide, a host cell comprising the
 CC expression cassette, and a transgenic plant comprising the expression
 CC cassette. The polypeptides and polynucleotides include those associated
 CC with the biosynthesis of isopentenyl diphosphate (IPP) and dimethylallyl
 CC alcohol (DMAPP), the biosynthesis of short-chain plastid
 CC prenyltransferases, the biosynthesis of gibberellins, the biosynthesis of
 CC carotenoids and/or abscisic acids, the biosynthesis of tocopherols,
 CC plastoquinone and/or phytylquinone biosynthesis, the mevalonate pathway,
 CC phytosterol and brassinosteroid metabolism, biosynthesis of ubiquinone,
 CC biosynthesis of monoterpenes and sesquiterpenes, protein prenylation, and
 CC biosynthesis of chlorophyll or haeme. Also disclosed are banana, wheat
 CC and corn homologues of some of the rice polynucleotides. The
 CC polynucleotides are useful for producing transgenic plants, where the
 CC genome is augmented by a nucleic acid molecule of the invention, or in
 CC which the corresponding gene has been disrupted, e.g. to result in a
 CC loss, a decrease or an alteration in the function of the product encoded
 CC by the gene. The plants may also have increased yields and/or produce a
 CC better quality product than the corresponding wild-type plant. The
 CC nucleic acid molecules are useful for targeted gene disruption, as well
 CC as markers and probes. Note: The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20040010815. The present sequence
 CC is an isoprenoid biosynthesis- associated DNA included in the sequence
 CC listing but not mentioned anywhere else in the specification.

SQ Sequence 657 BP; 109 A; 257 C; 178 G; 113 T; 0 U; 0 Other;

Query Match 16.9%; Score 30; DB 12; Length 657;
 Best Local Similarity 53.4%; Pred. No. 8.9;
 Matches 63; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 29 TGTGTTGCTAACCGCAGCGGACCACTTCCTATCGAACAGGACGGCGCTCCATATCA 88
 |||||
 DB 506 TGTGTGGCATGACACAGACGGACACAGACAGCTGTCTAAAGCGCGGACGGCC 447
 QY 89 GCCGCGCTTATCTCATCGGTGACCGGACACGAGCGCCCGTCCCGCTTATCGCGC 146
 |||||
 DB 446 GCCGCATGATTCGCTGTTCTGTGCGCGGTCTCGACGGGCGCGGTGCGCTCCGCGC 389
 |||||

RESULT 14

ABD14088/c
 ID ABD14088 standard; DNA; 1710 BP.

AC ABD14088;

XX 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polynucleotide #12692.

KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
 KW antibacterial.

OS Pseudomonas aeruginosa.

XX US6551795-B1.

PN 22-APR-2003.

PD 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

DR P-PSDB; ABO80517.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,

PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 12692; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABD01397-
 CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html

SQ Sequence 1710 BP; 380 A; 455 C; 526 G; 349 T; 0 U; 0 Other;

Query Match 16.9%; Score 30; DB 11; Length 1710;

Best Local Similarity 57.4%; Pred. No. 11;

Matches 54; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 81 CCATATCAGCGCGCGTATCTCATGCGGTGACGACGACGAGCGCGCTCCGCTTA 140
 |||||
 DB 414 CCAAAATAGCGCGTCTCATGCGGTGACGAGGACCTGCGAGCGCATTCGCCCC 355
 |||||

QY 141 TCGCGCTTATAATACAGCGCGCAACGATCTGGT 174
 |||||

DB 354 TCGCGCGCTCACTCTTCGATCCACCTTCGAGGT 321
 |||||

RESULT 15

ADS89385/c

ID ADS89385 standard; DNA; 3107 BP.

XX ADS89385;

XX 18-NOV-2004 (first entry)

DE Oligonucleotide of the invention SEQ ID NO:401.

KW ss; cell proliferative disorder; breast; methylation; cytostatic;
 KW gene therapy; single nucleotide polymorphism; SNP.

OS Unidentified.

XX WO2004035803-A2.

XX 29-APR-2004.

XX 01-OCT-2003; 2003WO-EP010881.

XX 01-OCT-2002; 2002DE-01045779.

PR 07-JAN-2003; 2003DE-01000096.

PR 17-APR-2003; 2003DE-01017955.

XX (EPIG-) EPIGENOMICS AG.

XX Foekens J, Harbeck N, Koenig T, Maier S, Martens J, Model P;

PI Nimmrich I, Rujan T, Schmitt A, Schmitt M, Look MP, Marx A;

XX WPI; 2004-348468/32.

XX Predicting responsiveness of a subject with breast cell proliferative
 PT disorder, useful for treating or differentiating breast cell
 PT proliferative disorders comprises analyzing methylation pattern of a
 PT genomic DNA from the subject.

XX Claim 25; SEQ ID NO 401; 104pp; English.
PS
XX
CC The invention relates to a novel method for predicting the responsiveness
CC of a subject with a cell proliferative disorder of the breast tissues to
CC a therapy comprising analysing the methylation pattern of a target
CC nucleic acid by contacting at least one of the target nucleic acids in a
CC biological sample obtained from the subject prior to or during treatment.
CC The method of the invention has cytosolic activity, and may have a use
CC in gene therapy. The set of oligonucleotides comprising at least two of
CC the oligomers are useful for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms (SNPs) within the sequences. The
CC methods, nucleic acid, oligonucleotide, and kit are useful for the
CC treatment, characterisation, classification and/or differentiation, of
CC breast cell proliferative disorders. The method is also useful for
CC predicting the responsiveness of a subject with a cell proliferative
CC disorder of the breast tissues to a therapy. The present sequence is used
CC in the exemplification of the invention.
XX
SQ Sequence 3107 BP; 823 A; 120 C; 814 G; 1350 T; 0 U; 0 Other;
Query Match 16.9%; Score 30; DB 13; Length 3107;
Best Local Similarity 50.7%; Pred. No. 13;
Matches 72; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 6 ATCGTGACAGGAGCGGCTTCCTGTTGCTAAACCGGAGCGGCACTCCTTATCG 65
Db 2079 AACGCGAAACACCCGCTCTCTCGCTCTCAACCGCGCGGAATAACCTCAACGCC 2020
QY 66 GAACAGGAGCGGCTTCATATACGCGCGGTTATCTCATGCGGTTGACCGGACAGG 125
Db 2019 GAACCGCGGACTTCTCTATTTAAACAACTTCTTAAACGCGCGGACAAACGACCCG 1960
QY 126 CGCCCGTCCGCTTATCGGCC 147
Db 1959 CACTTCTCCCAATATCGGCC 1938
Search completed: May 10, 2005, 04:32:48
Job time : 201.559 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 04:19:16 ; Search time 58.8508 Seconds
(without alignments)
4921.273 Million cell updates/sec

Title: US-09-896-888a-1_COPY_351_527

Perfect score: 177
Sequence: 1 gctttatcgtacagcgc.....gcccgaacgtctgtgtaa 177

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:..

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq*
- 5: /cgn2_6/ptodata/1/ina/PCFUS COMB.seq*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	30	16.9	1710	4	US-09-252-991A-12692
C 2	29.8	16.8	1311	4	US-09-902-540-2371
C 3	29.8	16.8	13579	4	US-09-902-540-1101
C 4	28.6	16.2	585	4	US-09-252-991A-12370
C 5	28.6	16.2	1443	4	US-09-252-991A-12097
C 6	28.6	16.2	1551	4	US-09-252-991A-12323
C 7	28.6	16.2	2085	4	US-09-252-991A-12237
C 8	28.6	16.2	2091	3	US-08-899-437-22
C 9	28.6	16.2	2091	3	US-09-126-121-22
C 10	28.6	16.2	2502	3	US-08-899-437-5
C 11	28.6	16.2	2502	3	US-09-126-121-5
C 12	28.6	16.2	25048	4	US-09-902-540-1239
C 13	28.4	16.0	57320	4	US-09-949-016-13983
C 14	28.2	15.9	283	4	US-09-513-9990-27806
C 15	28.2	15.9	2892	4	US-09-902-540-7217
C 16	28.2	15.9	3612	4	US-09-902-540-674
C 17	28.2	15.9	126200	4	US-09-949-016-11824
C 18	28.2	15.9	126200	4	US-09-949-016-13193
C 19	27.8	15.7	601	4	US-09-949-016-19425
C 20	27.8	15.7	601	4	US-09-949-016-49606
C 21	27.8	15.7	2208	4	US-09-902-540-7028
C 22	27.8	15.7	5588	4	US-09-902-540-621
C 23	27.6	15.6	534	4	US-09-489-039A-2740
C 24	27.6	15.6	798	4	US-09-489-039A-2763
C 25	27.4	15.5	37838	4	US-09-949-016-12134
C 26	27.2	15.4	549	4	US-09-902-540-6433
C 27	27.2	15.4	978	4	US-09-252-991A-14837

ALIGNMENTS

RESULT 1

US-09-252-991A-12692/c
; Sequence 12692, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12692
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12692

Query Match 16.9%; Score 30; DB 4; Length 1710;
Best Local Similarity 57.4%; Pred. No. 2;
Matches 54; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY	81	CCATATCAGCGCGCGTATCTCATCGCGGTACCGGACACGAGGCGCGTCCCGCTTA	140
Db	414	CCAAATACCGCTGATCTTCGGCGAAGTACGSGGACCTGGGAGGCGATTGCCCC	355
QY	141	TCGCGCTATAATACAGCCGCAACGATCTGGT	174
Db	354	TCGCGCGCTCAACTCTCTGCATCCACCTGCAGGT	321

RESULT 2

US-09-902-540-2371
; Sequence 2371, Application US/0902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10

Sequence 14970, A
Sequence 2474, Ap
Sequence 5551, Ap
Sequence 2099, Ap
Sequence 5635, Ap
Sequence 3318, Ap
Sequence 1963, Ap
Sequence 1048, Ap
Sequence 1110, Ap
Sequence 1154, Ap
Sequence 1269, Ap
Sequence 200, App
Sequence 1, Appl1
Sequence 193, App
Sequence 219, App
Sequence 228, App
Sequence 15207, A
Sequence 17417, A

```
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 2371
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-2371

Query Match      16.8%; Score 29.8; DB 4; Length 1311;
Best Local Similarity 55.2%; Pred. No. 2.1; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 47;

QY 45 GCCGACGCAACTCCTTATCGGAACAGACGCGCTCCATATCAGCCGCGGTTATCTCA 104
    |||||
Db 368 GCCGCACTTCTCGTCTGTGGACCGAGACACCTCCATCTTCTGAGCGGTGTCRACC 427
    |||||

QY 105 TCGCGGTACCGGACAGAGGCGCGCTCCGCTTATCGCGCCTA 149
    |||||
Db 428 GGAACGGGACATTGGACGCGCGCGGTCAAGCTCTTTACCCCCA 472
    |||||

RESULT 3
US-09-902-540-1101/c
; Sequence 1101, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1101
; LENGTH: 13579
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1101

Query Match      16.8%; Score 29.8; DB 4; Length 13579;
Best Local Similarity 55.2%; Pred. No. 4.3; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 47;

QY 45 GCCGACGCAACTCCTTATCGGAACAGACGCGCTCCATATCAGCCGCGGTTATCTCA 104
    |||||
Db 1063 GCCGCACTTCTCGTCTGTGGACCGAGACACCTCCATCTTCTGACGCGGTGTCRACC 1004
    |||||

QY 105 TCGCGGTACCGGACAGAGGCGCGCTCCGCTTATCGCGCCTA 149
    |||||
Db 1003 GGAACGGGACATTGGACGCGCGCGGTCAAGCTCTTTACCCCCA 959
    |||||

RESULT 4
US-09-252-991A-12370/c
; Sequence 12370, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12370
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12370
```

```
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12370

Query Match      16.2%; Score 28.6; DB 4; Length 585;
Best Local Similarity 53.0%; Pred. No. 4.3; Indels 0; Gaps 0;
Matches 61; Conservative 0; Mismatches 54;

QY 8 CGTGACAGGACGCGCAGCTTCTGTGTGTTGCTTAACCGGACCGGACCACTCTTATCGGA 67
    |||||
Db 534 CGTCCCTGTGGCAGGGGCTGTGTGTCGAGCGCGCTTCGATACCGGTGTCGA 475
    |||||

QY 68 ACAGACGCGGCTCCATATCAGCGCGCTTATCTCATGCGCGTACCGGACACG 122
    |||||
Db 474 GCTGAGCGCGGACAGATCAGCGGACGCTGGGGACGCTCGCGGAAATG 420
    |||||

RESULT 5
US-09-252-991A-12097
; Sequence 12097, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12097
; LENGTH: 1443
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12097

Query Match      16.2%; Score 28.6; DB 4; Length 1443;
Best Local Similarity 53.0%; Pred. No. 5.7; Indels 0; Gaps 0;
Matches 61; Conservative 0; Mismatches 54;

QY 8 CGTGACAGGACGCGCAGCTTCTGTGTGTTGCTTAACCGGACCGGACCACTCTTATCGGA 67
    |||||
Db 810 CGTCCCTGTGGCAGGGGCTGTGTGTCGAGCGCGCTTCGATACCGGTGTCGA 869
    |||||

QY 68 ACAGACGCGGCTCCATATCAGCGCGCTTATCTCATGCGCGTACCGGACACG 122
    |||||
Db 870 GCTGAGCGCGGACAGATCAAGCGGACGCTGGGGACGCTCGCGGAAATG 924
    |||||

RESULT 6
US-09-252-991A-12323/c
; Sequence 12323, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12323
; LENGTH: 1551
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12323
```

Query Match 16.2%; Score 28.6; DB 4; Length 1551;
Best Local Similarity 53.0%; Pred. No. 5.8;
Matches 61; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 8 CGTGACAGGAGCGCTTCTGTTGCTAAACGGAGCGGAGCAACTCTTATCGGA 67
DB 715 CGTGCCTGCTGGCAGGGGTGTGTGGAGCGCGATGCCGCTTCGATACCGTGTGCGA 656

QY 68 ACAGGAGCGCTTCCATATCAGCGCGGCTTATCTCATGCGGTGACCGGACACG 122
DB 655 GCTGGAGCGCGGAGATCAAGCGGAGGTGAGTGGGACGTCGCGCGGAATG 601

RESULT 7
US-09-252-991A-12237
; Sequence 12237, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12237
; LENGTH: 2085
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12237

Query Match 16.2%; Score 28.6; DB 4; Length 2085;
Best Local Similarity 53.0%; Pred. No. 6.3;
Matches 61; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 8 CGTGACAGGAGCGCTTCTGTTGCTAAACGGAGCGGAGCAACTCTTATCGGA 67
DB 802 CGTGCCTGCTGGCAGGGGTGTGTGGAGCGCGATGCCGCTTCGATACCGTGTGCGA 861

QY 68 ACAGGAGCGCTTCCATATCAGCGCGGCTTATCTCATGCGGTGACCGGACACG 122
DB 862 GCTGGAGCGCGGAGATCAAGCGGAGGTGAGTGGGACGTCGCGCGGAATG 916

RESULT 8
US-08-899-437-22
; Sequence 22, Application US/08899437
; Patent No. 6121415
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/08/899,437
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12237
; LENGTH: 2085
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-08-899-437-22

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Deirdre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 2091 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
FEATURE:
NAME/KEY: Human NRG3B2 (hNRGB2)
LOCATION: 1-2091
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-08-899-437-22

Query Match 16.2%; Score 28.6; DB 3; Length 2091;
Best Local Similarity 58.0%; Pred. No. 6.3;
Matches 69; Conservative 0; Mismatches 49; Indels 1; Gaps 1;

QY 55 ACTCTTATCGGAACAGGAGCGCTTCCATATCAGCGCGGCTTATCTCATGCGGTGAC 114
DB 448 ACGCCCAACCGGATTAGCACTCGCTGACCACTCAACGCGGCGCCCACTCGCTTCCC 506

QY 115 CGGACACAGGAGCGCGCTTATCGCGCTTATTAATACAGCGCGCAACGATCTGG 173
DB 507 CGGGACACCGGCTGCCCATCCGCGGCGCTCCACACAGCAGGAACTCGGG 565

RESULT 9
US-09-126-121-22
; Sequence 22, Application US/09126121
; Patent No. 6252051
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/126,121
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12237
; LENGTH: 2085
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-126-121-22

Query Match 16.2%; Score 28.6; DB 3; Length 2091;
Best Local Similarity 58.0%; Pred. No. 6.3;
Matches 69; Conservative 0; Mismatches 49; Indels 1; Gaps 1;

QY 55 ACTCTTATCGGAACAGGAGCGCTTCCATATCAGCGCGGCTTATCTCATGCGGTGAC 114
DB 448 ACGCCCAACCGGATTAGCACTCGCTGACCACTCAACGCGGCGCCCACTCGCTTCCC 506

QY 115 CGGACACAGGAGCGCGCTTATCGCGCTTATTAATACAGCGCGCAACGATCTGG 173
DB 507 CGGGACACCGGCTGCCCATCCGCGGCGCTCCACACAGCAGGAACTCGGG 565

```

; NAME/KEY: Human NR3B2 (hNRGB2)
; LOCATION: 1-2091
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
US-09-126-121-22
Query Match 16.2%; Score 28.6; DB 3; Length 2091;
Best Local Similarity 58.0%; Pred. No. 6.3;
Matches 69; Conservative 0; Mismatches 49; Indels 1; Gaps 1;
QY 55 ACTCTTATCGGAACAGAGCGCGCTCCATATCAGCGCGGTATCTATCTATCGCGGTGAC 114
Db 448 ACGCCCAACGGATTAGCACTCGCTGACCACCA-TCACGGGGGGCCCACTCGCTTCCC 506
QY 115 CGGACACGAGGCGCCGCTTATCGCGCTTATAATACAGCCCGCAACGATCTGG 173
Db 507 CGGGACCGGTGCCCATCCGGCCAGCCCGGCTCCACACGGAACACTGCGG 565
RESULT 10:
US-08-899-437-5
; Sequence 5, Application US/08899437
; Patent No. 6121415
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 23
; FILING DATE: 24-Jul-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Conley, Deirdre L.
; REGISTRATION NUMBER: 36,487
; REFERENCE/DOCKET NUMBER: P1084R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/952-9881
; TELEFAX: 650/225-2066
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,437
; FILING DATE: 24-Jul-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Conley, Deirdre L.
; REGISTRATION NUMBER: 36,487
; REFERENCE/DOCKET NUMBER: P1084R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/952-9881
; TELEFAX: 650/225-2066
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2502 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; FEATURE:
; NAME/KEY: Human NR3B1 (hNRGB1)/nucleic acid seq.
; LOCATION: 1-2502
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
US-08-899-437-5
Query Match 16.2%; Score 28.6; DB 3; Length 2502;
Best Local Similarity 58.0%; Pred. No. 6.7;
Matches 69; Conservative 0; Mismatches 49; Indels 1; Gaps 1;
QY 55 ACTCTTATCGGAACAGAGCGCGCTCCATATCAGCGCGGTATCTATCTATCGCGGTGAC 114
Db 591 ACGCCCAACGGATTAGCACTCGCTGACCACCA-TCACGGGGGGCCCACTCGCTTCCC 649
QY 115 CGGACACGAGGCGCCGCTTATCGCGCTTATAATACAGCCCGCAACGATCTGG 173
Db 507 CGGGACCGGTGCCCATCCGGCCAGCCCGGCTCCACACGGAACACTGCGG 565
RESULT 11
US-09-126-121-5
; Sequence 5, Application US/09126121
; Patent No. 6252051
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/126,121
; FILING DATE: 30-Jul-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Conley, Deirdre L.
; REGISTRATION NUMBER: 36,487
; REFERENCE/DOCKET NUMBER: P1084R1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-2066
; TELEFAX: 650/952-9881
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/126,121
; FILING DATE: 30-Jul-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Conley, Deirdre L.
; REGISTRATION NUMBER: 36,487
; REFERENCE/DOCKET NUMBER: P1084R1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-2066
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2502 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; FEATURE:
; NAME/KEY: Human NR3B1 (hNRGB1)/nucleic acid seq.
; LOCATION: 1-2502
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
US-09-126-121-5
Query Match 16.2%; Score 28.6; DB 3; Length 2502;
Best Local Similarity 58.0%; Pred. No. 6.7;
Matches 69; Conservative 0; Mismatches 49; Indels 1; Gaps 1;
QY 55 ACTCTTATCGGAACAGAGCGCGCTCCATATCAGCGCGGTATCTATCTATCGCGGTGAC 114
Db 591 ACGCCCAACGGATTAGCACTCGCTGACCACCA-TCACGGGGGGCCCACTCGCTTCCC 649
QY 115 CGGACACGAGGCGCCGCTTATCGCGCTTATAATACAGCCCGCAACGATCTGG 173
Db 507 CGGGACCGGTGCCCATCCGGCCAGCCCGGCTCCACACGGAACACTGCGG 708
RESULT 12
US-09-902-540-1239
; Sequence 1239, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wisigand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
```

```

; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1239
; LENGTH: 25048
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; US-09-902-540-1239

Query Match      16.2%; Score 28.6; DB 4; Length 25048;
Best Local Similarity 49.0%; Pred. No. 13;
Matches 76; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 4 TTATCGTGCACGAGCGCGCTTCTGTGTGCTTAACCGCAGCGGACGCAACTCCTTAT 63
Db 6260 TTTTGTGACGACACTCACCCATCGATGTTGGTACAGGGGGTGGAAACATCGAAATCA 6319

Qy 64 CGGACAGGAGCGCGCTTCATATCAGCCGCGGGTTATCTCATGCGCGGTGACCGGACAGA 123
Db 6320 CGTTTCATGAGGTTCCCGCGGAAGCAGCCCTTTGGCAACTCTCACATTTGGGATCGTGAGCGT 6379

Qy 124 GCGCCCGTCCCGCTTATCGCGCCTATAAATACAG 158
Db 6380 GCGCGCGTCCCGCACCCTTCAGAGTAAAGACAG 6414

RESULT 13
US-09-949-016-13983/c
; Sequence 13983 Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13983
; LENGTH: 57320
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(57320)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13983

Query Match      16.0%; Score 28.4; DB 4; Length 57320;
Best Local Similarity 49.3%; Pred. No. 20;
Matches 74; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 15 GGAGCGCAGCTTCTGTGTTGCTTAACCGCAGCGGACGCAACTCCTTATCGGAACAGGAC 74
Db 2034 GGAGCGCAGCTTCTGTGAGCGCGGACGCGCGCGGACGCGACGCAACCGTTC 1975

Qy 75 GCGGCTTCATATCAGCGCGGGTTATCTCATGCGGTGACCGGACACGAGGCGCGCGTCC 134
Db 1974 GCGGCTCACTTCGCGGAGCGCGCGCCCTGTCAACGCGTCAACGCGGACTCCGCCCGCGG 1915

Qy 135 GCCTTATCGCGCTTATAAATACAGCCCGGA 164
Db 1914 GCGCGGGGCGACCGGAGGACGAGCGGCA 1885

```

Db 1657 CAGGTGACGGGCTTCGGGGTTGCTCGCTCCAGCCCAAGCCGAAGCCAGGGGAGTAC 1716
Oy 73 ACGCGCCTCCATATCAGCGCGCGTTATCTCATGCGCGTGACCCG 117
Db 1717 GCGGCGCTGCTGTCATCGGCGGTACCAACAGAGCCGCGGCCAG 1761

Search completed: May 10, 2005, 07:29:35
Job time : 60.8508 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 05:52:55 ; Search time 500.12 Seconds
(without alignments)
2164.037 Million cell updates/sec

Title: US-09-896-888A-1_COPY_351_527

Perfect score: 177
Sequence: 1 gcttcatgtgacagagc.....gccgcaacgtctgtgtaa 177

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 5654200 seqs, 3057283753 residues

Total number of hits satisfying chosen parameters: 11308400

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA.*

- 1: /cgn2_6/ptodata1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata1/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata1/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata1/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	177	100.0	462	9	US-09-896-888A-14
2	177	100.0	560	18	US-10-622-088-126
3	177	100.0	584	9	US-09-896-888A-1
4	177	100.0	2773	16	US-10-295-074-60
5	177	100.0	2773	18	US-10-846-911-60
6	175.4	99.1	5038	18	US-10-622-088-89
7	41	23.2	147	18	US-10-622-088-127
8	41	23.2	325	18	US-10-622-088-149
9	32.2	18.2	3138	15	US-10-156-761-2538
10	32.2	18.2	9025608	15	US-10-156-761-1
11	30.6	17.3	1404	17	US-10-369-493-45503

c	12	30.2	17.1	1116	18	US-10-437-963-42126	Sequence 42126, A
c	13	30.2	17.1	2472	18	US-10-437-963-55473	Sequence 55473, A
c	14	30	16.9	657	17	US-10-259-194A-563	Sequence 563, App
c	15	30	16.9	6107	17	US-10-221-613-303	Sequence 303, App
c	16	29.8	16.8	1370	17	US-10-275-311A-10	Sequence 10, Appl
c	17	29.6	16.7	1798	18	US-10-739-930-4094	Sequence 4094, Ap
c	18	29.6	16.7	3498	18	US-10-437-963-72869	Sequence 72869, A
c	19	29.4	16.6	1923	18	US-10-437-963-50761	Sequence 50761, A
c	20	29.4	16.5	5118	18	US-10-437-963-35867	Sequence 35867, A
c	21	29.2	16.5	1593	17	US-10-369-493-44133	Sequence 44133, A
c	22	28.6	16.2	648	18	US-10-425-115-65984	Sequence 65984, A
c	23	28.6	16.2	823	16	US-10-029-386-22976	Sequence 22976, A
c	24	28.6	16.2	1425	9	US-09-815-242-4072	Sequence 4072, Ap
c	25	28.6	16.2	1425	17	US-10-282-122A-7326	Sequence 7326, Ap
c	26	28.6	16.2	1434	17	US-10-369-493-32534	Sequence 32534, A
c	27	28.6	16.2	2091	9	US-09-817-647-22	Sequence 22, Appl
c	28	28.6	16.2	2091	9	US-09-877-665-22	Sequence 22, Appl
c	29	28.6	16.2	2091	13	US-10-136-573A-22	Sequence 22, Appl
c	30	28.6	16.2	2091	14	US-10-215-862-22	Sequence 22, Appl
c	31	28.6	16.2	2091	19	US-10-944-116-22	Sequence 22, Appl
c	32	28.6	16.2	2502	9	US-09-817-647-5	Sequence 5, Appli
c	33	28.6	16.2	2502	9	US-09-877-665-5	Sequence 5, Appli
c	34	28.6	16.2	2502	13	US-10-136-573A-5	Sequence 5, Appli
c	35	28.6	16.2	2502	14	US-10-215-862-5	Sequence 5, Appli
c	36	28.6	16.2	2502	19	US-10-944-116-5	Sequence 5, Appli
c	37	28.4	16.0	876	17	US-10-369-493-41603	Sequence 41603, A
c	38	28.4	16.0	1320	17	US-10-282-122A-33339	Sequence 33339, A
c	39	28.4	16.0	1494	17	US-10-425-114-17146	Sequence 17146, A
c	40	28.4	16.0	2019	18	US-10-425-115-176225	Sequence 176225, A
c	41	28.4	16.0	4627	18	US-10-425-115-40821	Sequence 40821, A
c	42	28.4	16.0	29322	17	US-10-034-650-58	Sequence 58, Appl
c	43	28.4	16.0	75839	13	US-10-087-192-166	Sequence 166, App
c	44	28.2	15.9	1011	18	US-10-425-115-135441	Sequence 135441, A
c	45	28.2	15.9	1623	17	US-10-275-311A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-896-888A-14
; Sequence 14, Application US/09896888A
; Patent No. US20020116723A1
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/896,888A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/048,911
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/049,946
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 14
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Promoter
; OTHER INFORMATION: sequence of the OpMNPV ie2 gene
US-09-896-888A-14

Query Match	100.0%;	Score 177;	DB 9;	Length 462;
Best Local Similarity	100.0%;	Pred. No. 1.5e-54;		
Mismatches	177;	Conservative	0;	Mismatches
Indels	0;	Gaps	0;	
QY	1	GTCTTATCTGTGACAGGACGCCAGCTTCCCTGTGTGTTGCTTAACCGCAGCGCGACGCACTCCT	60	
Db	250	GTCTTATCTGTGACAGGACGCCAGCTTCCCTGTGTGTTGCTTAACCGCAGCGCGACGCACTCCT	309	
QY	61	TATCGGAACAGGACGCGCCTCCATATCAGCCGCGCGTTATCTCTATGCGGTGACCGGACA	120	

Db 310 TATCGAAGACGAGCGCTCCATATCAGCGCGGTTATCTCATGCGGTGACCGGACA 369
Qy 121 CGAGCGCGCGTCCCGCTTATCGCGCTTATAAATACAGCCGCGACGATCTGGTAAA 177
Db 370 CGAGCGCGCGTCCCGCTTATCGCGCTTATAAATACAGCCGCGACGATCTGGTAAA 426

RESULT 2

US-10-622-088-126
; Sequence 126, Application US/10622088
; Publication No. US20040219516A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 126
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: OplE2 promoter sequence
US-10-622-088-126

Query Match 100.0%; Score 177; DB 18; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.5e-54;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTCTTATCGTGACAGCGCCAGCTTCCTGTGTTGCTTAACCGCAGCGCGCAACTCCT 60
Db 355 GTCTTATCGTGACAGCGCCAGCTTCCTGTGTTGCTTAACCGCAGCGCGCAACTCCT 414
Qy 61 TATCGGAACAGGACGCGCTCCATATCAGCGCGGTTATCTCATGCGGTGACCGGACA 120
Db 415 TATCGGAACAGGACGCGCTCCATATCAGCGCGGTTATCTCATGCGGTGACCGGACA 474
Qy 121 CGAGCGCGCGTCCCGCTTATCGCGCTTATAAATACAGCCGCGACGATCTGGTAAA 177
Db 475 CGAGCGCGCGTCCCGCTTATCGCGCTTATAAATACAGCCGCGACGATCTGGTAAA 531

RESULT 3

US-09-896-888a-1
; Sequence 1, Application US/09896888A
; Patent No. US20020116723A1
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/896,888A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/048,911

; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/049,946
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Orgyia pseudotsugata
; US-09-896-888A-1

Query Match 100.0%; Score 177; DB 9; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.5e-54;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTCTTATCGTGACAGGACGCGCTTCCTGTGTTGCTTAACCGCAGCGCGCAACTCCT 60
Db 351 GTCTTATCGTGACAGGACGCGCTTCCTGTGTTGCTTAACCGCAGCGCGCAACTCCT 410
Qy 61 TATCGGAACAGGACGCGCTCCATATCAGCGCGGTTATCTCATGCGGTGACCGGACA 120
Db 411 TATCGGAACAGGACGCGCTCCATATCAGCGCGGTTATCTCATGCGGTGACCGGACA 470
Qy 121 CGAGCGCGCGTCCCGCTTATCGCGCTTATAAATACAGCCGCGACGATCTGGTAAA 177
Db 471 CGAGCGCGCGTCCCGCTTATCGCGCTTATAAATACAGCCGCGACGATCTGGTAAA 527

RESULT 4

US-10-295-074-60
; Sequence 60, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: P1013DK00
; CURRENT APPLICATION NUMBER: US/10/295,074
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 2773
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p2Op2F expression vector for insect cells
; FEATURE:
; NAME/KEY: misc recomb
; LOCATION: (561)..(566)
; OTHER INFORMATION: HindIII site
; FEATURE:
; NAME/KEY: misc recomb
; LOCATION: (573)..(578)
; OTHER INFORMATION: AvaI site
; FEATURE:
; NAME/KEY: misc recomb
; LOCATION: (586)..(591)
; OTHER INFORMATION: EcoRI site
; FEATURE:
; NAME/KEY: misc recomb
; LOCATION: (593)..(598)
; OTHER INFORMATION: BamHI site
; FEATURE:
; NAME/KEY: misc recomb
; LOCATION: (625)..(630)
; OTHER INFORMATION: ClaI site
; FEATURE:
; NAME/KEY: misc recomb
; LOCATION: (629)..(634)
; OTHER INFORMATION: ClaI site
; FEATURE:
; NAME/KEY: misc recomb
; LOCATION: (1156)..(1161)

```
; OTHER INFORMATION: ApaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2128)..(2133)
; OTHER INFORMATION: PstI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2204)..(2209)
; OTHER INFORMATION: NcoI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2284)..(2289)
; OTHER INFORMATION: AvalI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2294)..(2299)
; OTHER INFORMATION: AvalI, SmaI, and XmaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2551)..(2556)
; OTHER INFORMATION: ApaI site
US-10-295-074-60

Query Match      100.0%; Score 177; DB 16; Length 2773;
Best Local Similarity 100.0%; Pred. No. 1.9e-54;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCTTATCGTGACAGGACGCGCTTCCCTGTTGTTGCTTAACCGCAGCGCGGACGCAACTCCT 60
Db 355 GTCTTATCGTGACAGGACGCGCTTCCCTGTTGTTGCTTAACCGCAGCGCGGACGCAACTCCT 414

QY 61 TATCGGAACAGGACGCGCTTCCATATACAGCGCGCTTATCTCATGCGGTGACCGGACA 120
Db 415 TATCGGAACAGGACGCGCTTCCATATACAGCGCGCTTATCTCATGCGGTGACCGGACA 474

QY 121 CGAGGCGCGCTCCCGCTTATCGGCTTATAAATACAGCGCGCAACGATCTGGTAAA 177
Db 475 CGAGGCGCGCTCCCGCTTATCGGCTTATAAATACAGCGCGCAACGATCTGGTAAA 531

RESULT 5
US-10-846-911-60
; Sequence 60, Application US/10846911
; Publication No. US20040258660A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen
; APPLICANT: NIELSEN, Finn Stausholm
; APPLICANT: BRATT, Tomas
; APPLICANT: VOLDORF, Bjorn
; APPLICANT: MOURITSEN, Soren
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: 674542-2018
; CURRENT APPLICATION NUMBER: US/10/846,911
; PRIOR FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: PCT/DK02/00764
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: 60/331,575
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: PA 2001 01702
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 2773
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p2Op2F expression vector for insect cells
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (561)..(566)
; OTHER INFORMATION: HindIII site
; FEATURE:
```

```
; NAME/KEY: misc_recomb
; LOCATION: (573)..(578)
; OTHER INFORMATION: AvalI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (586)..(591)
; OTHER INFORMATION: EcoRI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (593)..(598)
; OTHER INFORMATION: BamHI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (625)..(630)
; OTHER INFORMATION: ClaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (629)..(634)
; OTHER INFORMATION: ClaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (1156)..(1161)
; OTHER INFORMATION: ApaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2128)..(2133)
; OTHER INFORMATION: PstI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2204)..(2209)
; OTHER INFORMATION: NcoI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2284)..(2289)
; OTHER INFORMATION: AvalI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2294)..(2299)
; OTHER INFORMATION: ApaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2551)..(2556)
; OTHER INFORMATION: ApaI site
US-10-846-911-60

Query Match      100.0%; Score 177; DB 18; Length 2773;
Best Local Similarity 100.0%; Pred. No. 1.9e-54;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCTTATCGTGACAGGACGCGCTTCCCTGTTGTTGCTTAACCGCAGCGCGGACGCAACTCCT 60
Db 355 GTCTTATCGTGACAGGACGCGCTTCCCTGTTGTTGCTTAACCGCAGCGCGGACGCAACTCCT 414

QY 61 TATCGGAACAGGACGCGCTTCCATATACAGCGCGCTTATCTCATGCGGTGACCGGACA 120
Db 415 TATCGGAACAGGACGCGCTTCCATATACAGCGCGCTTATCTCATGCGGTGACCGGACA 474

QY 121 CGAGGCGCGCTCCCGCTTATCGGCTTATAAATACAGCGCGCAACGATCTGGTAAA 177
Db 475 CGAGGCGCGCTCCCGCTTATCGGCTTATAAATACAGCGCGCAACGATCTGGTAAA 531

RESULT 6
US-10-622-088-89
; Sequence 89, Application US/10622088
; Publication No. US20040219516A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth
; APPLICANT: Franke, Kenneth E.
```

```
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 89
; LENGTH: 5038
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pIB/V5-His-DEST
US-10-622-088-89

Query Match          99.1%; Score 175.4; DB 18; Length 5038;
Best Local Similarity 99.4%; Pred. No. 8.2e-54;
Matches 176; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  GTCATTATCGTGACAGAGCCGAGCTTCCTGTTGCTTAACCGCAGCCGCGCAACTCCT 60
Db      351 GTCATTATCGTGACAGAGCCGAGCTTCCTGTTGCTTAACCGCAGCCGCGCAACTCCT 410

Qy      61  TATCGGAACAGAGCCGCGCTCATATACGCGCGGTTTATCTCATGCGGTGACCGGACA 120
Db      411 TATCGGAACAGAGCCGCGCTCATATACGCGCGGTTTATCTCATGCGGTGACCGGACA 470

Qy      121 CGAGCGCCCGTCCGCTTATCGGCTTATAATACAGCCGCGCAACTCTGGTAAA 177
Db      471 CGAGCGCCCGTCCGCTTATCGGCTTATAATACAGCCGCGCAACTCTGGTAAA 527

RESULT 7
US-10-622-088-127
; Sequence 127, Application US/10622088
; Publication No. US20040219516A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 149
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Recombination region of pIB/V5 His DEST
; NAME/KEY: misc feature
; LOCATION: (141)..(142)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (145)..(276)
US-10-622-088-149

Query Match          23.2%; Score 41; DB 18; Length 325;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      137 CTTATCGCGCTATAAATACAGCCCGCAACGATCTGGTAAA 177
Db      1  CTTATCGCGCTATAAATACAGCCCGCAACGATCTGGTAAA 41

RESULT 8
US-10-622-088-149
; Sequence 149, Application US/10622088
; Publication No. US20040219516A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 149
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Recombination region of pIB/V5 His DEST
; NAME/KEY: misc feature
; LOCATION: (141)..(142)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (145)..(276)
US-10-622-088-149

Query Match          23.2%; Score 41; DB 18; Length 325;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      137 CTTATCGCGCTATAAATACAGCCCGCAACGATCTGGTAAA 177
Db      1  CTTATCGCGCTATAAATACAGCCCGCAACGATCTGGTAAA 41
```

```
; LENGTH: 147
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombination region of pIB/V5-His-DEST
; NAME/KEY: misc feature
; LOCATION: (141)..(148)
; OTHER INFORMATION: n may be any nucleotide
US-10-622-088-127

Query Match          23.2%; Score 41; DB 18; Length 147;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      137 CTTATCGCGCTATAAATACAGCCCGCAACGATCTGGTAAA 177
Db      1  CTTATCGCGCTATAAATACAGCCCGCAACGATCTGGTAAA 41

RESULT 8
US-10-622-088-149
; Sequence 149, Application US/10622088
; Publication No. US20040219516A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 149
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Recombination region of pIB/V5 His DEST
; NAME/KEY: misc feature
; LOCATION: (141)..(142)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (145)..(276)
US-10-622-088-149

Query Match          23.2%; Score 41; DB 18; Length 325;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      137 CTTATCGCGCTATAAATACAGCCCGCAACGATCTGGTAAA 177
Db      1  CTTATCGCGCTATAAATACAGCCCGCAACGATCTGGTAAA 41
```

RESULT 9

US-10-156-761-2538/c
; Sequence 2538, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 2538
; LENGTH: 3138
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3138)
US-10-156-761-2538

Query Match 18.2%; Score 32.2; DB 15; Length 3138;
Best Local Similarity 61.2%; Pred. No. 0.22;
Matches 52; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 58 CCTTATCGAACAGACGGCGCTCCATATCAGCGCGCTTATCTCATCGCGGTGACCGG 117
DB 2575 CCTCACCAGAACCGACGGCGCTCCCTTGATGAGGTGAGTTGCTTCGCACGACCGG 2516

QY 118 ACACGAGGCGCGCTCCCTTATC 142
DB 2515 CCTTGAGGCGCTCGTCCAGGTTCTC 2491

RESULT 10

US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 18.2%; Score 32.2; DB 15; Length 9025608;
Best Local Similarity 61.2%; Pred. No. 0.78;
Matches 52; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 58 CCTTATCGAACAGACGGCGCTCCATATCAGCGCGCTTATCTCATCGCGGTGACCGG 117
DB 3129487 CCTCACCAGAACCGACGGCGCTCCCTTGATGAGGTGAGTTGCTTCGCACGACCGG 3129546

QY 118 ACACGAGGCGCGCTCCCTTATC 142
DB 3129547 CCTTGAGGCGCTCGTCCAGGTTCTC 3129571

RESULT 11

US-10-369-493-45503
; Sequence 45503, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 45503
; LENGTH: 1404
; TYPE: DNA
; ORGANISM: Rhodobacter capsulatus
US-10-369-493-45503

Query Match 17.3%; Score 30.6; DB 17; Length 1404;
Best Local Similarity 62.3%; Pred. No. 0.74;
Matches 48; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 99 ATCTCATGCGGTGACCGGACGAGGCGCGCTCCGCTTATCGCGCTTATCGGCTTAATAATACAG 158
DB 464 ACCGCAAGGCGGTGAACGTCGCCGATGTGCTGCTGCGGCTTCGCGGCTTCGCGCAAGG 523

QY 159 CCCGCAACGATCTGGTA 175
DB 524 ACCGCGACGATCTGGAA 540

RESULT 12

US-10-437-963-42126/c
; Sequence 42126, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Boukharov, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 42126
; LENGTH: 1116
; TYPE: DNA

```
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_45406C.1
US-10-437-963-42126

Query Match      17.1%; Score 30.2; DB 18; Length 1116;
Best Local Similarity 53.4%; Pred. No. 1.1;
Matches 62; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 16 GACGCCAGCTTCCTGTGTGTAAACCGGAGCGGACGCAACTCTTATCGGAACAGGACG 75
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 188 GCCGCCGCGTGTGTGTGTAAACCGGAGCGGAGCGGACGCAACTCTTATCGGACGATCTCCG 129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 76 CGCTCCATATACGCGCGCTTATCTCATGCGGTGACCGGACGAGGCGGCC 130
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128 CCCCCCGCTCTCTCCCGACCGGATCCGACGCGGCGCTCGGTGACGTGAGGCC 74
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-10-437-963-55473
; Sequence 55473, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437.963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 55473
; LENGTH: 2472
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2472)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_57478C.1
US-10-437-963-55473

Query Match      17.1%; Score 30.2; DB 18; Length 2472;
Best Local Similarity 55.1%; Pred. No. 1.1;
Matches 59; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 38 AACCGCAGCGGACGCAACTCTTATCGGAACAGGACGCGCTTCCATATCAGCGCGCGT 97
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1543 AACGTCATCGACGACGCGTTCGCTTCCCTCCGACGCGGCGGCGGCGGCGTTCGCGGT 1602
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 98 TATCTCATGCGGTGACCGGACGAGGCGGCGGCTCCGCTTATCGC 144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1603 GACCTCATCGGCGGCTGCTGCTCAAGGAGGCCCAACAGCGGATCGC 1649
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
US-10-259-194A-563/c
; Sequence 563, Application US/10259194A
; Publication No. US20040010815A1
; GENERAL INFORMATION:
; APPLICANT: Lange, Markus B.
; APPLICANT: Ghassenian, Majid
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
```

```
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kieps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF PLANT GENES
; FILE REFERENCE: 70029-NP
; CURRENT APPLICATION NUMBER: US/10/259,194A
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,743
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 662
; SOFTWARE: Patentlist.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 563
; LENGTH: 657
; TYPE: DNA
; ORGANISM: Triticum aestivum
; US-10-259-194A-563

Query Match      16.9%; Score 30; DB 17; Length 657;
Best Local Similarity 53.4%; Pred. No. 1.1;
Matches 63; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 29 TGTGTTCTTAACCGACGCGGACGCAACTCTTATCGGAACAGGACGCGCTCCATATCA 88
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 506 TGTGTGTCATGTCACACGACGCGGACGCAACTCTTATCGGAACAGGACGCGCTCCATATCA 447
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 89 GCGCGCGTTATCTCATGCGGTGACCGGACGAGGCGGCGGCTTATCGCGC 146
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 446 GCGCATGATGCTGTTCTGTCGCGGTCTGACGCGGCGGCTGCGCTCCGCGC 389
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
US-10-221-613-303/c
; Sequence 303, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 303
; LENGTH: 6107
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-303

Query Match      16.9%; Score 30; DB 17; Length 6107;
Best Local Similarity 50.7%; Pred. No. 1.6;
Matches 72; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
```


5

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 04:33:07 ; Search time 1465.23 Seconds
(without alignments)
4931.434 Million cell updates/sec

Title: US-09-896-888A-1_COPY_351_527

Perfect score: 177

Sequence: 1 gctctatctgcagagcgc.....gcccgaacgacgtgtgtaaa 177

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_NA_Main:*

- 1: /cgn2_6/ptodata/1/pna/PCTUS1_COMB.seq:*
- 2: /cgn2_6/ptodata/1/pna/PCTUS2_COMB.seq:*
- 3: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq:*
- 4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
- 5: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
- 6: /cgn2_6/ptodata/1/pna/US08_COMB.seq:*
- 7: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
- 8: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
- 9: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
- 10: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
- 11: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
- 12: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
- 13: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
- 14: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
- 15: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
- 16: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
- 17: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
- 18: /cgn2_6/ptodata/1/pna/US092A_COMB.seq:*
- 19: /cgn2_6/ptodata/1/pna/US092B_COMB.seq:*
- 20: /cgn2_6/ptodata/1/pna/US093A_COMB.seq:*
- 21: /cgn2_6/ptodata/1/pna/US093B_COMB.seq:*
- 22: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
- 23: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
- 24: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
- 25: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
- 26: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
- 27: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
- 28: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
- 29: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
- 30: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
- 31: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
- 32: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
- 33: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
- 34: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
- 35: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
- 36: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
- 37: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
- 38: /cgn2_6/ptodata/1/pna/US098D_COMB.seq:*
- 39: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
- 40: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
- 41: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
- 42: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
- 43: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*

44: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:*

45: /cgn2_6/ptodata/1/pna/US099G_COMB.seq:*

46: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*

47: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*

48: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*

49: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*

50: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*

51: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*

52: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*

53: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*

54: /cgn2_6/ptodata/1/pna/US104A_COMB.seq:*

55: /cgn2_6/ptodata/1/pna/US104B_COMB.seq:*

56: /cgn2_6/ptodata/1/pna/US105A_COMB.seq:*

57: /cgn2_6/ptodata/1/pna/US105B_COMB.seq:*

58: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*

59: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*

60: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*

61: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*

62: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*

63: /cgn2_6/ptodata/1/pna/US108A_COMB.seq:*

64: /cgn2_6/ptodata/1/pna/US108B_COMB.seq:*

65: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*

66: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*

67: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*

68: /cgn2_6/ptodata/1/pna/US110_COMB.seq:*

69: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*

70: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*

71: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*

72: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*

73: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*

74: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*

75: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*

76: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*

77: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*

78: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*

79: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*

80: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*

81: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*

82: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*

83: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*

84: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*

85: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*

86: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*

87: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*

88: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*

89: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*

90: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*

91: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*

92: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq:*

93: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq:*

94: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*

95: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*

96: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*

97: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*

98: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*

99: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*

100: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*

101: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*

102: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*

103: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*

104: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*

105: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*

106: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*

107: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*

108: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*

109: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*

110: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*

111: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*

112: /cgn2_6/ptodata/1/pna/US6042_COMB.seq:*

113: /cgn2_6/ptodata/1/pna/US6043_COMB.seq:*

114: /cgn2_6/ptodata/1/pna/US6044_COMB.seq:*

115: /cgn2_6/ptodata/1/pna/US6045_COMB.seq:*

116: /cgn2_6/ptodata/1/pna/US6046_COMB.seq:*

117: /cgn2_6/prodata/1/pna/US6047_COMB.seq.*
 118: /cgn2_6/prodata/1/pna/US6048_COMB.seq.*
 119: /cgn2_6/prodata/1/pna/US6049_COMB.seq.*
 120: /cgn2_6/prodata/1/pna/US6050_COMB.seq.*
 121: /cgn2_6/prodata/1/pna/US6051_COMB.seq.*
 122: /cgn2_6/prodata/1/pna/US6052_COMB.seq.*
 123: /cgn2_6/prodata/1/pna/US6053_COMB.seq.*
 124: /cgn2_6/prodata/1/pna/US6054_COMB.seq.*
 125: /cgn2_6/prodata/1/pna/US6055_COMB.seq.*
 126: /cgn2_6/prodata/1/pna/US6056_COMB.seq.*
 127: /cgn2_6/prodata/1/pna/US6057_COMB.seq.*
 128: /cgn2_6/prodata/1/pna/US6058_COMB.seq.*
 129: /cgn2_6/prodata/1/pna/US6059_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	177	100.0	462	16	US-09-048-911-14	Sequence 14, Appl
2	177	100.0	462	38	US-09-896-888-14	Sequence 14, Appl
3	177	100.0	462	38	US-09-896-888A-14	Sequence 14, Appl
4	177	100.0	560	2	PCT-US03-22437-126	Sequence 126, Appl
5	177	100.0	560	58	US-10-622-088-126	Sequence 126, Appl
6	177	100.0	564	16	US-09-048-911-1	Sequence 1, Appl
7	177	100.0	564	38	US-09-896-888-1	Sequence 1, Appl
8	177	100.0	564	38	US-09-896-888A-1	Sequence 1, Appl
9	177	100.0	2773	51	US-10-295-074-60	Sequence 60, Appl
10	177	100.0	2773	63	US-10-846-911-60	Sequence 60, Appl
11	175.4	99.1	5038	2	PCT-US03-22437-89	Sequence 89, Appl
12	175.4	99.1	5038	58	US-10-622-088-89	Sequence 89, Appl
13	41	23.2	147	2	PCT-US03-22437-127	Sequence 127, Appl
14	41	23.2	147	58	US-10-622-088-127	Sequence 127, Appl
15	41	23.2	325	58	US-10-622-088-149	Sequence 149, Appl
16	34	19.2	22899	32	US-09-702-134-16542	Sequence 16542, A
17	34	19.2	22899	35	US-09-815-264-72544	Sequence 72544, A
18	34	19.2	23229	28	US-09-620-392-920	Sequence 920, Appl
19	33.4	18.9	492	19	US-09-293-972-29963	Sequence 29963, A
20	33.4	18.9	492	39	US-09-904-939-29963	Sequence 29963, A
21	32.4	18.3	1194	46	US-10-015-127-5182	Sequence 5182, Ap
22	32.4	18.3	25201	46	US-10-015-127-514	Sequence 514, Appl
23	32.4	18.3	48610	28	US-09-620-392-22894	Sequence 22894, A
24	32.4	18.3	48610	32	US-09-702-134-27502	Sequence 27502, A
25	32.4	18.3	48610	35	US-09-815-264-81375	Sequence 81375, A
26	32.2	18.2	3138	49	US-10-156-761-2538	Sequence 2538, Ap
27	32.2	18.2	9025608	49	US-10-156-761-1	Sequence 1, Appl
28	32	18.1	921	20	US-09-313-292-2465	Sequence 2465, Ap
29	32	18.1	921	62	US-10-779-543-8561	Sequence 8561, Ap
30	31.4	17.7	418	27	US-09-606-977-59093	Sequence 59093, A
31	31.4	17.7	418	83	US-60-141-233-59093	Sequence 59093, A
32	31.4	17.7	425	36	US-09-837-604B-43377	Sequence 43377, A
33	31.4	17.7	425	36	US-09-837-604B-43377	Sequence 43377, A
34	31.4	17.7	425	88	US-60-197-872-40261	Sequence 40261, A
35	30.8	17.4	1268	27	US-09-614-150-25442	Sequence 25442, A
36	30.8	17.4	1268	27	US-09-614-150A-25442	Sequence 25442, A
37	30.8	17.4	1268	86	US-60-173-464-21087	Sequence 21087, A
38	30.8	17.4	1268	88	US-60-191-637-25561	Sequence 25561, A
39	30.8	17.4	1268	88	US-60-191-681-20170	Sequence 20170, A
40	30.8	17.4	1342	27	US-09-614-150-38495	Sequence 38495, A
41	30.8	17.4	1342	27	US-09-614-150A-38495	Sequence 38495, A
42	30.8	17.4	1342	88	US-60-191-637-38121	Sequence 38121, A
43	30.8	17.4	3448	27	US-09-614-150-38494	Sequence 38494, A
44	30.8	17.4	3448	27	US-09-614-150A-38494	Sequence 38494, A
45	30.8	17.4	3448	88	US-60-191-637-38120	Sequence 38120, A

ALIGNMENTS

RESULT 1
 US-09-048-911-14
 ; Sequence 14, Application US/09048911
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of British Columbia
 ; TITLE OF INVENTION: Insect Expression Vectors
 ; FILE REFERENCE: 80021-44
 ; CURRENT APPLICATION NUMBER: US/09/048,911
 ; CURRENT FILING DATE: 1998-03-26
 ; EARLIER APPLICATION NUMBER: US 60/049,946
 ; EARLIER FILING DATE: 1997-03-27
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 14
 ; LENGTH: 462
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Promoter
 ; OTHER INFORMATION: sequence of the OpMPV iez gene
 US-09-048-911-14
 Query Match 100.0%; Score 177; DB 16; Length 462;
 Best Local Similarity 100.0%; Pred. No. 1.3e-44;
 Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GTCTTATCGTGACAGGAGCGCGCTTCTCTGTTGCTAAACCGCAGCGCGCAACTCCT 60
 |||||
 Db 250 GTCTTATCGTGACAGGAGCGCGCTTCTCTGTTGCTAAACCGCAGCGCGCAACTCCT 309
 |||||
 Qy 61 TATCGGAACAGGAGCGCGCTTCTATATCAGCGCGCGTATCTCTATGCGCGTGACCGGACA 120
 |||||
 Db 310 TATCGGAACAGGAGCGCGCTTCTATATCAGCGCGCGTATCTCTATGCGCGTGACCGGACA 369
 |||||
 Qy 121 CGAGCGCGCGTCCCGCTTATCGCGCTTATAATACAGCGCGCAACGATCTGGTAAA 177
 |||||
 Db 370 CGAGCGCGCGTCCCGCTTATCGCGCTTATAATACAGCGCGCAACGATCTGGTAAA 426
 |||||
 RESULT 2
 US-09-896-888-14
 ; Sequence 14, Application US/09896888
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of British Columbia
 ; TITLE OF INVENTION: Insect Expression Vectors
 ; FILE REFERENCE: 80021-44
 ; CURRENT APPLICATION NUMBER: US/09/896,888
 ; CURRENT FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: US 09/048,911
 ; PRIOR FILING DATE: 1998-03-26
 ; PRIOR APPLICATION NUMBER: US 60/049,946
 ; PRIOR FILING DATE: 1997-03-27
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 14
 ; LENGTH: 462
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Promoter
 ; OTHER INFORMATION: sequence of the OpMPV iez gene
 US-09-896-888-14
 Query Match 100.0%; Score 177; DB 38; Length 462;
 Best Local Similarity 100.0%; Pred. No. 1.3e-44;
 Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GTCTTATCGTGACAGGAGCGCGCTTCTCTGTTGCTAAACCGCAGCGCGCAACTCCT 60
 |||||
 Db 250 GTCTTATCGTGACAGGAGCGCGCTTCTCTGTTGCTAAACCGCAGCGCGCAACTCCT 309
 |||||
 Qy 61 TATCGGAACAGGAGCGCGCTTCTATATCAGCGCGCGTATCTCTATGCGCGTGACCGGACA 120
 |||||

```
Db 310 TATCGGAACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 369
QY 121 CGAGGCGCGCGTCCCGCTTATCGCGCTATAAATACAGCCGCAACGATCTGTGTA 177
Db 370 CGAGGCGCGCGTCCCGCTTATCGCGCTATAAATACAGCCGCAACGATCTGTGTA 426

RESULT 3
US-09-896-888A-14
; Sequence 14, Application US/09896888A
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/896,888A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/048,911
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/049,946
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Promoter
; OTHER INFORMATION: sequence of the OpMNPV ie2 gene
US-09-896-888A-14

Query Match 100.0%; Score 177; DB 38; Length 462;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCCTATCGTACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 369
Db 250 GTCCTATCGTACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 309
QY 61 TATCGGAACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 120
Db 310 TATCGGAACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 369
QY 121 CGAGGCGCGCGTCCCGCTTATCGCGCTATAAATACAGCCGCAACGATCTGTGTA 177
Db 370 CGAGGCGCGCGTCCCGCTTATCGCGCTATAAATACAGCCGCAACGATCTGTGTA 426

RESULT 4
PCT-US03-22437-126
; Sequence 126, Application PC/TUS0322437
; GENERAL INFORMATION:
; APPLICANT: Invitrogen Corporation
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.545PC07
; CURRENT APPLICATION NUMBER: PCT/US03/22437
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 126
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Opie2 promoter sequence
US-10-622-088-126

Query Match 100.0%; Score 177; DB 58; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCCTATCGTACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 60
Db 355 GTCCTATCGTACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 414
QY 61 TATCGGAACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 120
Db 415 TATCGGAACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 474
QY 121 CGAGGCGCGCGTCCCGCTTATCGCGCTATAAATACAGCCGCAACGATCTGTGTA 177
Db 475 CGAGGCGCGCGTCCCGCTTATCGCGCTATAAATACAGCCGCAACGATCTGTGTA 531
```

```
; FEATURE:
; OTHER INFORMATION: Opie2 promoter sequence
PCT-US03-22437-126

Query Match 100.0%; Score 177; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCCTATCGTACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 60
Db 355 GTCCTATCGTACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 414
QY 61 TATCGGAACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 120
Db 415 TATCGGAACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 474
QY 121 CGAGGCGCGCGTCCCGCTTATCGCGCTATAAATACAGCCGCAACGATCTGTGTA 177
Db 475 CGAGGCGCGCGTCCCGCTTATCGCGCTATAAATACAGCCGCAACGATCTGTGTA 531

RESULT 5
US-10-622-088-126
; Sequence 126, Application US/10622088
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 126
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Opie2 promoter sequence
US-10-622-088-126

Query Match 100.0%; Score 177; DB 58; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCCTATCGTACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 60
Db 355 GTCCTATCGTACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 414
QY 61 TATCGGAACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 120
Db 415 TATCGGAACAGGAGCGCCTCCATATACAGCCGCGCTTATCTCATCGCGGTGACCGGACA 474
QY 121 CGAGGCGCGCGTCCCGCTTATCGCGCTATAAATACAGCCGCAACGATCTGTGTA 177
Db 475 CGAGGCGCGCGTCCCGCTTATCGCGCTATAAATACAGCCGCAACGATCTGTGTA 531
```

```
RESULT 6
US-09-048-911-1
; Sequence 1, Application US/09048911
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/048,911
; CURRENT FILING DATE: 1998-03-26
; EARLIER APPLICATION NUMBER: US 60/049,946
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; TYPE: DNA
; LENGTH: 564
; ORGANISM: Orgyia pseudotsugata
US-09-048-911-1

Query Match      100.0%; Score 177; DB 16; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GTCTTATCGTGACAGAGCGCCAGCTTCTCTGTGTTGCTAAACCGCAGCGCGGACGCAACTCCT 60
DB      351  GTCTTATCGTGACAGAGCGCCAGCTTCTCTGTGTTGCTAAACCGCAGCGCGGACGCAACTCCT 410

QY      61  TATCGGAACAGGACGCGCTCCATATCAGCGCGCGGTTATCTCATGCGCGTGACCGGACA 120
DB      411  TATCGGAACAGGACGCGCTCCATATCAGCGCGCGGTTATCTCATGCGCGTGACCGGACA 470

QY      121  CGAGGCGCGCGTCCCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCTGGTAAA 177
DB      471  CGAGGCGCGCGTCCCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCTGGTAAA 527

RESULT 7
US-09-896-888-1
; Sequence 1, Application US/09896888
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/896,888
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/048,911
; PRIOR FILING DATE: 1998-03-26
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; TYPE: DNA
; LENGTH: 564
; ORGANISM: Orgyia pseudotsugata
US-09-896-888-1

Query Match      100.0%; Score 177; DB 38; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GTCTTATCGTGACAGAGCGCCAGCTTCTCTGTGTTGCTAAACCGCAGCGCGGACGCAACTCCT 60
DB      351  GTCTTATCGTGACAGAGCGCCAGCTTCTCTGTGTTGCTAAACCGCAGCGCGGACGCAACTCCT 410

QY      61  TATCGGAACAGGACGCGCTCCATATCAGCGCGCGGTTATCTCATGCGCGTGACCGGACA 120
DB      411  TATCGGAACAGGACGCGCTCCATATCAGCGCGCGGTTATCTCATGCGCGTGACCGGACA 470

QY      121  CGAGGCGCGCGTCCCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCTGGTAAA 177
DB      471  CGAGGCGCGCGTCCCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCTGGTAAA 527

RESULT 8
US-09-896-888A-1
; Sequence 1, Application US/09896888A
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/896,888A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/048,911
; PRIOR FILING DATE: 1998-03-26
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; TYPE: DNA
; LENGTH: 564
; ORGANISM: Orgyia pseudotsugata
US-09-896-888A-1

Query Match      100.0%; Score 177; DB 38; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GTCTTATCGTGACAGAGCGCCAGCTTCTCTGTGTTGCTAAACCGCAGCGCGGACGCAACTCCT 60
DB      351  GTCTTATCGTGACAGAGCGCCAGCTTCTCTGTGTTGCTAAACCGCAGCGCGGACGCAACTCCT 410

QY      61  TATCGGAACAGGACGCGCTCCATATCAGCGCGCGGTTATCTCATGCGCGTGACCGGACA 120
DB      411  TATCGGAACAGGACGCGCTCCATATCAGCGCGCGGTTATCTCATGCGCGTGACCGGACA 470

QY      121  CGAGGCGCGCGTCCCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCTGGTAAA 177
DB      471  CGAGGCGCGCGTCCCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCTGGTAAA 527

RESULT 9
US-10-295-074-60
; Sequence 60, Application US/10295074
; GENERAL INFORMATION:
; APPLICANT: Phatwexa A/S
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: P1013DK00
; CURRENT APPLICATION NUMBER: US/10/295,074
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 2773
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p2Op2F expression vector for insect cells
; NAME/KEY: misc_recomb
; LOCATION: (561)..(566)
; OTHER INFORMATION: HindIII site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (573)..(578)
; OTHER INFORMATION: Aval site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (586)..(591)
; OTHER INFORMATION: EcoRI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (593)..(598)
; OTHER INFORMATION: BamHI site
```

FEATURE:
NAME/KEY: misc_recomb
LOCATION: (625)..(630)
OTHER INFORMATION: ClaI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (629)..(634)
OTHER INFORMATION: ClaI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (1156)..(1161)
OTHER INFORMATION: ApaLI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2128)..(2133)
OTHER INFORMATION: PstI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2204)..(2209)
OTHER INFORMATION: NcoI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2284)..(2289)
OTHER INFORMATION: AvalI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2294)..(2299)
OTHER INFORMATION: AvaI, SmaI, and XmaI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2551)..(2556)
OTHER INFORMATION: ApaLI site
US-10-295-074-60

Query Match 100.0%; Score 177; DB 51; Length 2773;
Best Local Similarity 100.0%; Pred. No. 1.6e-44;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCTTATCGTGACAGAGCGCCAGCTTCTCTGTGTGCTTAACCGCAGCGCGGCAACTCTCT 60
DB 355 GTCTTATCGTGACAGAGCGCCAGCTTCTCTGTGTGCTTAACCGCAGCGCGGCAACTCTCT 414
QY 61 TATCGGAACAGAGCGCGCTTCCATATCAGCGCGCGCTTATCTCATCGCGGTGACCGGACA 120
DB 415 TATCGGAACAGAGCGCGCTTCCATATCAGCGCGCGCTTATCTCATCGCGGTGACCGGACA 474
QY 121 CGAGGCGCGCGTCCCGCTTATCGCGCTTATAATACAGCGCGCAACGATCTGTATAA 177
DB 475 CGAGGCGCGCGTCCCGCTTATCGCGCTTATAATACAGCGCGCAACGATCTGTATAA 531

RESULT 10
US-10-846-911-60
; Sequence 60, Application US/10846911
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen
; APPLICANT: NIELSEN, Finn Stausholm
; APPLICANT: BRATT, Tomas
; APPLICANT: VOLDORF, Bjorn
; APPLICANT: MOURITSEN, Soren
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: 674542-2018
; CURRENT APPLICATION NUMBER: US/10/846, 911
; CURRENT FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: PCT/DK02/00764
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: 60/331,575
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: PA 2001 01702
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60

LENGTH: 2773
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: p2Op2F expression vector for insect cells
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (561)..(566)
OTHER INFORMATION: HindIII site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (573)..(578)
OTHER INFORMATION: AvalI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (586)..(591)
OTHER INFORMATION: EcoRI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (593)..(598)
OTHER INFORMATION: BamHI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (625)..(630)
OTHER INFORMATION: ClaI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (629)..(634)
OTHER INFORMATION: ClaI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (1156)..(1161)
OTHER INFORMATION: ApaLI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2128)..(2133)
OTHER INFORMATION: PstI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2204)..(2209)
OTHER INFORMATION: NcoI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2284)..(2289)
OTHER INFORMATION: AvalI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2294)..(2299)
OTHER INFORMATION: AvaI, SmaI, and XmaI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2551)..(2556)
OTHER INFORMATION: ApaLI site
US-10-846-911-60

Query Match 100.0%; Score 177; DB 63; Length 2773;
Best Local Similarity 100.0%; Pred. No. 1.6e-44;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCTTATCGTGACAGAGCGCCAGCTTCTCTGTGTGCTTAACCGCAGCGCGCAACTCTCT 60
DB 355 GTCTTATCGTGACAGAGCGCCAGCTTCTCTGTGTGCTTAACCGCAGCGCGCAACTCTCT 414
QY 61 TATCGGAACAGAGCGCGCTTCCATATCAGCGCGCGCTTATCTCATCGCGGTGACCGGACA 120
DB 415 TATCGGAACAGAGCGCGCTTCCATATCAGCGCGCGCTTATCTCATCGCGGTGACCGGACA 474
QY 121 CGAGGCGCGCGTCCCGCTTATCGCGCTTATAATACAGCGCGCAACGATCTGTATAA 177
DB 475 CGAGGCGCGCGTCCCGCTTATCGCGCTTATAATACAGCGCGCAACGATCTGTATAA 531
RESULT 11

```
PCT-US03-22437-89
; Sequence 89, Application PC/TUS0322437
; GENERAL INFORMATION:
; APPLICANT: Invitrogen Corporation
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.545PC07
; CURRENT APPLICATION NUMBER: PCT/US03/22437
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 89
; LENGTH: 5038
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pIB/V5-His-DEST
PCT-US03-22437-89

Query Match          99.1%; Score 175.4; DB 2; Length 5038;
Best Local Similarity 99.4%; Pred. No. 5.5e-44;
Matches 176; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTCATTATCGTGACAGGAGCGCGCTTCCATATACAGCGCGGTTATCTCATGCGCGTGACCGGACA 60
Db 351 GTCATTATCGTGACAGGAGCGCGCTTCCATATACAGCGCGGTTATCTCATGCGCGTGACCGGACA 410

Qy 61 TATCGGAACAGGAGCGCGCTTCCATATACAGCGCGGTTATCTCATGCGCGTGACCGGACA 120
Db 411 TATCGGAACAGGAGCGCGCTTCCATATACAGCGCGGTTATCTCATGCGCGTGACCGGACA 470

Qy 121 CGAGCGCGCGTCCGCTTATCGCGCTTATATAATACAGCGCGCAACGATCTGGTAAA 177
Db 471 CGAGCGCGCGTCCGCTTATCGCGCTTATATAATACAGCGCGCAACGATCTGGTAAA 527

RESULT 12
US-10-622-088-89
; Sequence 89, Application US/10622088
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 89
; LENGTH: 5038
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pIB/V5-His-DEST
US-10-622-088-89

Query Match          99.1%; Score 175.4; DB 58; Length 5038;
Best Local Similarity 99.4%; Pred. No. 5.5e-44;
Matches 176; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTCATTATCGTGACAGGAGCGCGCTTCCATATACAGCGCGGTTATCTCATGCGCGTGACCGGACA 60
Db 351 GTCATTATCGTGACAGGAGCGCGCTTCCATATACAGCGCGGTTATCTCATGCGCGTGACCGGACA 410

Qy 61 TATCGGAACAGGAGCGCGCTTCCATATACAGCGCGGTTATCTCATGCGCGTGACCGGACA 120
Db 411 TATCGGAACAGGAGCGCGCTTCCATATACAGCGCGGTTATCTCATGCGCGTGACCGGACA 470

Qy 121 CGAGCGCGCGTCCGCTTATCGCGCTTATATAATACAGCGCGCAACGATCTGGTAAA 177
Db 471 CGAGCGCGCGTCCGCTTATCGCGCTTATATAATACAGCGCGCAACGATCTGGTAAA 527

RESULT 13
PCT-US03-22437-127
; Sequence 127, Application PC/TUS0322437
; GENERAL INFORMATION:
; APPLICANT: Invitrogen Corporation
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.545PC07
; CURRENT APPLICATION NUMBER: PCT/US03/22437
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 127
; LENGTH: 147
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombination region of pIB/V5-His-DEST
PCT-US03-22437-127

Query Match          23.2%; Score 41; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 137 CTTATCGCGCTATAAATACAGCGCGCAACGATCTGGTAAA 177
Db 1 CTTATCGCGCTATAAATACAGCGCGCAACGATCTGGTAAA 41

RESULT 14
US-10-622-088-127
; Sequence 127, Application US/10622088
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
```

```

; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Primpong, Kenneth E.
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 127
; LENGTH: 147
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Recombination region of pIB/V5-His-DEBT
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (141)..(148)
; OTHER INFORMATION: n may be any nucleotide
US-10-622-088-127

```

```

Query Match      23.2%; Score 41; DB 58; Length 147;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      137  CTTATCGCGCCTATAATAACAGCCCGCAACGATCTGGTAAA 177
Db      1    CTTATCGCGCCTATAATAACAGCCCGCAACGATCTGGTAAA 41

```

```

RESULT 15
US-10-622-088-149
; Sequence 149, Application US/10622088
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Primpong, Kenneth
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 149
; LENGTH: 325

```

```

; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Recombination region of pIB/V5 His DEBT
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (141)..(142)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (145)..(276)
US-10-622-088-149

Query Match      23.2%; Score 41; DB 58; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      137  CTTATCGCGCCTATAATAACAGCCCGCAACGATCTGGTAAA 177
Db      1    CTTATCGCGCCTATAATAACAGCCCGCAACGATCTGGTAAA 41

Search completed: May 10, 2005, 09:18:56
Job time : 1472.23 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 05:51:35 ; Search time 436.346 Seconds
(without alignments)

834.449 Million cell updates/sec

Title: US-09-896-888A-1_COPY_351_527

Perfect score: 177

Sequence: 1 gcttctatcgtacaggagcgc.....gcccgaacgatctgtgtaa 177

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 9398789 seqs, 102855566 residues

Total number of hits satisfying chosen parameters: 18797578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_NA_New.*
1: /cgn2_6/ptodata/1/pna/US06_PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_PCT_NEW_COMB.seq2.*
3: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq2.*
9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq3.*
10: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq2.*
12: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq3.*
13: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	177	100.0	2773	7	US-10-939-107-60
2	30.8	17.4	1268	10	US-11-097-143-25442
3	30.8	17.4	1342	10	US-11-097-143-25442
C 4	30.8	17.4	1348	10	US-11-097-143-38495
C 5	30.8	17.4	3448	10	US-11-097-143-38495
C 6	30.6	17.3	3600	10	US-11-097-143-38494
7	30.6	17.3	10968	2	PCT-US05-07924-35
8	30.6	17.3	10968	11	US-11-075-185-35
9	30.6	17.3	78869	11	PCT-US05-07924-1
C 10	30.6	17.3	78869	11	US-11-075-185-1
C 11	30	16.9	3107	8	US-10-517-441-401
C 12	29.8	16.8	3501	8	US-10-517-441-499
C 13	29.8	16.8	1311	11	US-11-031-175-2371
C 14	29.8	16.8	1789	10	US-11-097-143-26108
C 15	29.8	16.8	3915	10	US-11-097-143-26107
C 16	28.6	16.2	13579	11	US-11-031-175-1101
C 17	28.6	16.2	547	9	US-10-703-032-96240
C 18	28.6	16.2	551	9	US-10-703-032-96240
C 19	28.6	16.2	915	8	US-10-450-763-52910
C 20	28.6	16.2	1703	9	US-10-703-032-518
21	28.6	16.2	2091	11	US-11-035-787-22

21	28.6	16.2	2502	11	US-11-035-787-5	Sequence 5, Appli
22	28.6	16.2	25048	11	US-11-031-175-1239	Sequence 1239, Ap
C 23	28.6	16.2	215013	13	US-60-659-397-11844	Sequence 11844, A
C 24	28.2	15.9	2225	8	US-10-450-763-14117	Sequence 14117, A
C 25	28.2	15.9	2892	11	US-11-031-175-7217	Sequence 7217, Ap
C 26	28.2	15.9	3612	11	US-11-031-175-674	Sequence 674, App
C 27	27.8	15.7	500	9	US-10-703-032-82313	Sequence 82313, A
C 28	27.8	15.7	579	9	US-10-703-032-28994	Sequence 28994, A
C 29	27.8	15.7	1200	9	US-10-703-032-26385	Sequence 26385, A
C 30	27.8	15.7	2208	11	US-11-031-175-7028	Sequence 7028, Ap
C 31	27.8	15.7	5588	11	US-11-031-175-621	Sequence 621, App
C 32	27.4	15.5	601	9	US-10-703-032-82054	Sequence 82054, A
C 33	27.4	15.5	816	10	US-11-097-143-27146	Sequence 27146, A
C 34	27.4	15.5	1707	8	US-10-450-763-1430	Sequence 1430, Ap
C 35	27.4	15.5	5314	10	US-11-097-143-27145	Sequence 27145, A
C 36	27.4	15.5	5643	1	PCT-US05-13243-5	Sequence 5, Appli
C 37	27.4	15.5	5643	10	US-11-109-593-5	Sequence 5, Appli
C 38	27.4	15.5	5679	2	PCT-US05-07924-36	Sequence 36, Appli
C 39	27.4	15.5	5679	11	US-11-075-185-36	Sequence 36, Appli
C 40	27.4	15.5	67323	1	PCT-US05-13243-1	Sequence 1, Appli
C 41	27.4	15.5	67323	10	US-11-109-593-1	Sequence 1, Appli
C 42	27.4	15.5	188504	11	US-11-033-545-711	Sequence 711, App
C 43	27.4	15.5	200918	11	US-11-033-545-647	Sequence 647, App
C 44	27.2	15.4	549	11	US-11-031-175-6433	Sequence 6433, Ap
C 45	27.2	15.4	600	9	US-10-972-079-37529	Sequence 37529, A

ALIGNMENTS

RESULT 1

US-10-939-107-60
; Sequence 60, Application US/10939107
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Hans Rudolf
; APPLICANT: Ebert, Bjarke
; APPLICANT: Pedersen, Louise Henriette
; APPLICANT: Rasmussen, Peter Birk
; TITLE OF INVENTION: Novel Application of Vaccination Against TNF-alpha
; FILE REFERENCE: 674542-2020
; CURRENT APPLICATION NUMBER: US/10/939,107
; CURRENT FILING DATE: 2004-09-10
; PRIOR APPLICATION NUMBER: PCT/DK03/00147
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: 60/363,128
; PRIOR FILING DATE: 2002-03-11
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 60
; LENGTH: 2773
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p2Op2F expression vector for insect cells
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (561)..(566)
; OTHER INFORMATION: HindIII site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (573)..(578)
; OTHER INFORMATION: Aval site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (586)..(591)
; OTHER INFORMATION: EcoRI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (593)..(598)
; OTHER INFORMATION: BamHI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (625)..(630)

OTHER INFORMATION: ClaI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (629)..(634)
OTHER INFORMATION: ClaI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (1156)..(1161)
OTHER INFORMATION: ApaLI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2128)..(2133)
OTHER INFORMATION: PstI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2204)..(2209)
OTHER INFORMATION: NcoI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2284)..(2289)
OTHER INFORMATION: AvalI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2294)..(2299)
OTHER INFORMATION: AvalI, SmaI, and XmaI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2551)..(2556)
OTHER INFORMATION: ApaLI site
US-10-939-107-60

Query Match 100.0%; Score 177; DB 7; Length 2773;
Best Local Similarity 100.0%; Pred. No. 1e-44;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCTTATCGTACAGAGCGGCTTCTGTTGCTAACCAGCGGAGCGCAACTCCT 60
DB 355 GTCTTATCGTACAGAGCGGCTTCTGTTGCTAACCAGCGGAGCGCAACTCCT 414
QY 61 TATCGGACAGAGCGGCTTCTATATACAGCGGCTTATCTATGCGGTGACCGGACA 120
DB 415 TATCGGACAGAGCGGCTTCTATATACAGCGGCTTATCTATGCGGTGACCGGACA 474
QY 121 CGAGCGCGCGTCCGCTTATCGCGCTTATATAACAGCGGCAACGATCTGGTAAA 177
DB 475 CGAGCGCGCGTCCGCTTATCGCGCTTATATAACAGCGGCAACGATCTGGTAAA 531

RESULT 2
US-11-097-143-25442
Sequence 25442, Application US/11097143
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
APPLICANT: et al.
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
TITLE OF INVENTION: DROSOPHILA GENES.
FILE REFERENCE: CL000728
CURRENT APPLICATION NUMBER: US/11/097,143
CURRENT FILING DATE: 2005-04-04
PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/160,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/164,769
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/173,383
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: 60/175,693
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: 60/184,831
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637
PRIOR FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 43008
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 38495
LENGTH: 1342
TYPE: DNA
ORGANISM: DROSOPHILA
US-11-097-143-38495

Query Match 17.4%; Score 30.8; DB 10; Length 1342;
Best Local Similarity 55.7%; Pred. No. 4.8;
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 17 AGCCAGCTTCCTGTTGCTAACCAGCGGAGCGCAACTCCTTATCGGAACGAGGCGC 76
DB 731 ACAACAGCTTCCTGTTGAGGTGCCCAAGCGGCTTCTAGTCTTGGCGGCTCCGGTGGC 790
QY 77 GCCTCCATATCAGCGCGGCTTATCTATGCGGTGACCGGACG 122
DB 791 GTTGGACATGGCCATACAAATCTCAAGTCCAGAAAGTACAAG 836

PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637
PRIOR FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 43008
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25442
LENGTH: 1268
TYPE: DNA
ORGANISM: DROSOPHILA
US-11-097-143-25442

Query Match 17.4%; Score 30.8; DB 10; Length 1268;
Best Local Similarity 55.7%; Pred. No. 4.8;
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 17 AGCCAGCTTCCTGTTGCTAACCAGCGGAGCGCAACTCCTTATCGGAACGAGGCGC 76
DB 657 ACAACAGCTTCCTGTTGAGGTGCCCAAGCGGCTTCTAGTCTTGGCGGCTCCGGTGGC 716
QY 77 GCCTCCATATCAGCGCGGCTTATCTATGCGGTGACCGGACG 122
DB 717 GTTGGACATGGCCATACAAATCTCAAGTCCAGAAAGTACAAG 762

RESULT 3
US-11-097-143-38495
Sequence 38495, Application US/11097143
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
APPLICANT: et al.
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
TITLE OF INVENTION: DROSOPHILA GENES.
FILE REFERENCE: CL000728
CURRENT APPLICATION NUMBER: US/11/097,143
CURRENT FILING DATE: 2005-04-04
PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/160,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/164,769
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/173,383
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: 60/175,693
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: 60/184,831
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637
PRIOR FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 43008
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 38495
LENGTH: 1342
TYPE: DNA
ORGANISM: DROSOPHILA
US-11-097-143-38495

Query Match 17.4%; Score 30.8; DB 10; Length 1342;
Best Local Similarity 55.7%; Pred. No. 4.8;
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 17 AGCCAGCTTCCTGTTGCTAACCAGCGGAGCGCAACTCCTTATCGGAACGAGGCGC 76
DB 731 ACAACAGCTTCCTGTTGAGGTGCCCAAGCGGCTTCTAGTCTTGGCGGCTCCGGTGGC 790
QY 77 GCCTCCATATCAGCGCGGCTTATCTATGCGGTGACCGGACG 122
DB 791 GTTGGACATGGCCATACAAATCTCAAGTCCAGAAAGTACAAG 836

```

RESULT 4
US-11-097-143-38494/c
; Sequence 38494, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38494
; LENGTH: 3448
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-38494

Query Match 17.4%; Score 30.8; DB 10; Length 3448;
Best Local Similarity 55.7%; Pred. No. 5.2;
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 17 ACGCCAGCTTCCTGTGTGTTAACCAGCGGAGCGGACCACTCTTATCGGACAGGACGC 76
Db 1667 ACAACAGCTTCCTGTGTGTTAACCAGCGGAGCGGACCACTCTTATCGGACAGGACGC 76
QY 77 GCCTCCATATCAGCGCGCGGTATCTCATGCGCGTGACCGGACACG 122
Db 1607 GTTGGCACATTGCCATACAAATCTCAAGTCCAGAAAGTACAG 1562

RESULT 5
US-11-097-143-25441/c
; Sequence 25441, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12

```

; FILE REFERENCE: 010099.03
; CURRENT APPLICATION NUMBER: US/11/075,185
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/551,103
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US 60/568,290
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 35
; LENGTH: 10968
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-11-075-185-35

Query Match 17.3%; Score 30.6; DB 11; Length 10968;
Best Local Similarity 53.3%; Pred. No. 6.6;
Matches 88; Conservative 0; Mismatches 74; Indels 3; Gaps 1;
QY 3 CTTATCGTGACGAGCGCGCTTCTCTGTTGCTTAACCGCAGCGCGGACGCAACTCCTTA 62
Db 6775 CTGCTCTCTGTCGGTGGGACGAGCGCGGCTGCGGCCCGGCGGCGCTCGCC--- 6831
QY 63 TCGGAACGAGGCGCGCTCCATATCAGCGCGCGGTATCTCATGCGCGTGACCGGACACG 122
Db 6832 GAGCACCTGCGCGCCACCGGACGAGCGGTGCTCGAGCTCGCGCGGAGCTGGCCACG 6891
QY 123 AGGCGCGCGTCCCGCTTATCGGGCTATAAATACAGCCCGCAACG 167
Db 6892 ACGGCGACGACCTCGCCACGCGGCTCGCGTGGCGCTCGCGCG 6936

RESULT 8

PCT-US05-07924-1
; Sequence 1, Application PC/TUS0507924

; GENERAL INFORMATION:
; APPLICANT: REEVES, CHRISTOPHER D
; APPLICANT: JULIEN, BRYAN
; APPLICANT: REID, RALPH
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
; FILE REFERENCE: 010099.03
; CURRENT APPLICATION NUMBER: PCT/US05/07924
; CURRENT FILING DATE: 2005-03-17
; PRIOR APPLICATION NUMBER: US 60/551,103
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US 60/568,290
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 78869
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
PCT-US05-07924-1

Query Match 17.3%; Score 30.6; DB 2; Length 78869;
Best Local Similarity 53.3%; Pred. No. 7.7;
Matches 88; Conservative 0; Mismatches 74; Indels 3; Gaps 1;
QY 3 CTTATCGTGACGAGCGCGCTTCTCTGTTGCTTAACCGCAGCGCGGACGCAACTCCTTA 62
Db 22042 CTGCTCTCTGTCGGTGGGACGAGCGCGCTGCGGCCCGGCGGCGCTCGCC--- 22098
QY 63 TCGGAACGAGGCGCGCTCCATATCAGCGCGCGGTATCTCATGCGGTGACCGGACACG 122
Db 22099 GAGCACCTGCGCGGCCACCGGACGAGCGGTGCTCGAGCTGCGCGGAGCTGGCCACG 22158
QY 123 AGGCGCGCGTCCCGCTTATCGGGCTATAAATACAGCCCGCAACG 167
Db 22159 ACGGCGACGACCTCGCCACGCGGCTCGCGTGGCGCTCGCGCG 22203

RESULT 9

US-11-075-185-1
; Sequence 1, Application US/11075185
; GENERAL INFORMATION:
; APPLICANT: REEVES, CHRISTOPHER D
; APPLICANT: JULIEN, BRYAN
; APPLICANT: REID, RALPH
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
; FILE REFERENCE: 010099.03
; CURRENT APPLICATION NUMBER: US/11/075,185
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/551,103
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US 60/568,290
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 78869
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-11-075-185-1

Query Match 17.3%; Score 30.6; DB 11; Length 78869;
Best Local Similarity 53.3%; Pred. No. 7.7;
Matches 88; Conservative 0; Mismatches 74; Indels 3; Gaps 1;
QY 3 CTTATCGTGACGAGCGCGCTTCTCTGTTGCTTAACCGCAGCGCGGACGCAACTCCTTA 62
Db 22042 CTGCTCTCTGTCGGTGGGACGAGCGCGCTGCGGCCCGGCGGCGCTCGCC--- 22098
QY 63 TCGGAACGAGGCGCGCTCCATATCAGCGCGCGGTATCTCATGCGGTGACCGGACACG 122
Db 22099 GAGCACCTGCGCGGCCACCGGACGAGCGGTGCTCGAGCTCGCGCGGAGCTGGCCACG 22158
QY 123 AGGCGCGCGTCCCGCTTATCGGGCTATAAATACAGCCCGCAACG 167
Db 22159 ACGGCGACGACCTCGCCACGCGGCTCGCGTGGCGCTCGCGCG 22203

RESULT 10

US-10-517-441-401/c
; Sequence 401, Application US/10517441

; GENERAL INFORMATION:
; APPLICANT: FOEKENS, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almuth
; APPLICANT: HOFER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 401
; LENGTH: 3107
; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)

US-10-517-441-401

Query Match 16.9%; Score 30; DB 8; Length 3107;
Best Local Similarity 50.7%; Pred. No. 9;
Matches 72; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 6 ATCGTGACAGAGCCGAGCTTCCTGTTGCTAAACCGGAGCGGCAACTCCTTATCG 65
DB 2079 AACGGAAACACCCGCTCTCTCGCTCTCCAAACCGCGCGGAAATAACCTCAACGGCC 2020

QY 66 GAACAGACGCGCTCATATACGCGCGGTTATCTCATGCGGTGACGGAGACGAGG 125
DB 2019 GAACCGCGGACTTCTCTCATTTTAAACAACTTCTTAAACGCGCGGCAACAAACGACCGG 1960

QY 126 CGCCCGTCCGCTTATCGGCC 147
DB 1959 CACTTCTCCCAATATCGCGCC 1938

RESULT 11

US-10-517-441-499/C

; Sequence 499, Application US/10517441

; GENERAL INFORMATION:

; APPLICANT: FOEKENS, John

; APPLICANT: HARBEC, Nadia

; APPLICANT: KOENIG, Thomas

; APPLICANT: MAIER, Sabine

; APPLICANT: MARTENS, John

; APPLICANT: MODEL, Fabian

; APPLICANT: NIMMICH, Inko

; APPLICANT: RUJAN, Tamas

; APPLICANT: SCHMITT, Armin

; APPLICANT: SCHMITT, Manfred

; APPLICANT: LOOK, Maxime P.

; APPLICANT: MARX, Almuth

; APPLICANT: HOEFLER, Heinz

; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell

; TITLE OF INVENTION: proliferative disorders

; FILE REFERENCE: 47675-93

; CURRENT APPLICATION NUMBER: US/10/517,441

; PRIOR FILING DATE: 2004-12-11

; PRIOR APPLICATION NUMBER: PCT/EP2003/010881

; PRIOR FILING DATE: 2003-10-01

; PRIOR APPLICATION NUMBER: DE 10317955.0

; PRIOR FILING DATE: 2003-04-17

; PRIOR APPLICATION NUMBER: DE 10300096.8

; PRIOR FILING DATE: 2003-01-07

; PRIOR APPLICATION NUMBER: DE 10245779.4

; PRIOR FILING DATE: 2002-10-01

; NUMBER OF SEQ ID NOS: 2147

; SEQ ID NO 499

; LENGTH: 3501

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)

US-10-517-441-499

Query Match 16.9%; Score 30; DB 8; Length 3501;
Best Local Similarity 50.7%; Pred. No. 9.1;
Matches 72; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 6 ATCGTGACAGAGCCGAGCTTCCTGTTGCTAAACCGGAGCGGCAACTCCTTATCG 65
DB 1585 AACGGAAACACCCGCTCTCTCGCTCTCCAAACCGCGCGGAAATAACCTCAACGGCC 1526

QY 66 GAACAGACGCGCTCATATACGCGCGGTTATCTCATGCGGTGACGGAGACGAGG 125
DB 1525 GAACCGCGGACTTCTCTCATTTTAAACAACTTCTTAAACGCGCGGCAACAAACGACCGG 1466

QY 126 CGCCCGTCCGCTTATCGGCC 147
DB 1465 CACTTCTCCCAATATCGGCC 1444

RESULT 12

US-11-031-175-2371

; Sequence 2371, Application US/11031175

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/11/031,175

; CURRENT FILING DATE: 2005-01-08

; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 2371

; LENGTH: 1311

; TYPE: DNA

; ORGANISM: Myxococcus xanthus

US-11-031-175-2371

Query Match 16.8%; Score 29.8; DB 11; Length 1311;
Best Local Similarity 55.2%; Pred. No. 9.7;
Matches 58; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 45 GCCGAGCGCAACTCTTATCGGAACAGGACGCGCTCCATATACGCCGCGTTATCTCA 104
DB 368 GCCGCGAGTTCTCTCGTCTGACCGGAGACACCTCCATCTTCTGACGCGGTCAACC 427

QY 105 TCGGCTGACCGGACGAGGCGCGCTCCGCTTATCGCGCTA 149
DB 428 GGAACGGGACATTGGACGCGCGCGGTCTCTTTACCCCA 472

RESULT 13

US-11-097-143-26108

; Sequence 26108, Application US/11097143

; GENERAL INFORMATION:

; APPLICANT: Venter, J. Craig

; APPLICANT: et al.

; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID

; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE

; FILE REFERENCE: CL000728

; CURRENT APPLICATION NUMBER: US/11/097,143

; PRIOR FILING DATE: 2005-04-04

; PRIOR APPLICATION NUMBER: 60/157,832

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: 60/160,191

; PRIOR FILING DATE: 1999-10-19

; PRIOR APPLICATION NUMBER: 60/161,932

; PRIOR FILING DATE: 1999-10-28

; PRIOR APPLICATION NUMBER: 60/164,769

; PRIOR FILING DATE: 1999-11-12

; PRIOR APPLICATION NUMBER: 60/173,383

; PRIOR FILING DATE: 1999-12-28

; PRIOR APPLICATION NUMBER: 60/175,693

; PRIOR FILING DATE: 2000-01-12

; PRIOR APPLICATION NUMBER: 60/184,831

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: 60/191,637

; PRIOR FILING DATE: 2000-03-23

; NUMBER OF SEQ ID NOS: 43008

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 26108

; LENGTH: 1789

; TYPE: DNA

; ORGANISM: DROSOPHILA

US-11-097-143-26108

Query Match 16.8%; Score 29.8; DB 10; Length 1789;
Best Local Similarity 49.1%; Pred. No. 9.9;
Matches 79; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 7 TCGTGACAGGACGGCAGCTTCTGTGTGCTTAACCGCAGCGCGAGCGCAACTCCTTATCGG 66
Db 1226 TCGTGCCAACTCTGTGTGCAACAGTGTCCAGTGACGAGCGACCCCATCTTGTGTGC 1285

Qy 67 AACAGGACGGCCTTCATATACAGCGCGCGTTATCTCATGCGGTGACCGGACACGAGGC 126
Db 1286 CACTGCTGGGACCCCTTCAACAACCAACAGGTGTTCATACAGCAGCAGCAACAA 1345

Qy 127 GCCGTCGCCGTTATCGCGCTTAATAATACAGCGCGCAACG 167
Db 1346 GCCCAAGTGTCTGTGCTGCCCTTGAACCAACGACGCGAAG 1386

RESULT 14

US-11-097-143-26107/c
; Sequence 26107, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26107
; LENGTH: 3915
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-26107

Query Match 16.8%; Score 29.8; DB 10; Length 3915;
Best Local Similarity 49.1%; Pred. No. 11;
Matches 79; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 7 TCGTGACAGGACGGCAGCTTCTGTGTGCTTAACCGCAGCGCGAGCGCAACTCCTTATCGG 66
Db 1564 TCGTGCCAACTCTGTGTGCAACAGTGTCCAGTGACGAGCGACCCCATCTTGTGTGC 1505

Qy 67 AACAGGACGGCCTTCATATACAGCGCGCGTTATCTCATGCGGTGACCGGACACGAGGC 126
Db 1504 CACTGCTGGGACCCCTTCAACAACCAACAGGTGTTCATACAGCAGCAGCAACAA 1445

Qy 127 GCCGTCGCCGTTATCGCGCTTAATAATACAGCGCGCAACG 167
Db 1444 GCCCAAGTGTCTGTGCTGCCCTTGAACCAAGCAGCGAAG 1404

RESULT 15

US-11-031-175-1101/c
; Sequence 1101, Application US/11031175
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/11/031,175
; CURRENT FILING DATE: 2005-01-08
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1101
; LENGTH: 13579
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-11-031-175-1101

Query Match 16.8%; Score 29.8; DB 11; Length 13579;
Best Local Similarity 55.2%; Pred. No. 12;
Matches 58; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 45 GCCGACGCAACTCTTATCGGAACAGGACGCGCTCCATATCAGCGCGGTTATCTCA 104
Db 1063 GCCGCGAGTTCCTGTCGTCTGGACCGGAGACACTTCATCTTCTGACGGGTGTCAC 1004

Qy 105 TGCGCGTGACCGGACACGAGCGCGCTCCGCTTATCGGCGCTTA 149
Db 1003 GGAACGGGACATTGGAGCGCGCGCGGTACGCTCTTTACCCCA 959

Search completed: May 10, 2005, 09:51:26
Job time : 437.346 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 03:45:00 ; Search time 1240.12 Seconds
(without alignments)

5432.847 Million cell updates/sec

Title: US-09-896-888a-1_COPY_351_527

Perfect score: 177

Sequence: 1 gcttatcgtagacagagc.....gccgcgaagatctggtgaaa 177

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33.8	19.1	507	8	AQ297217 HS 2266 A
2	33.6	19.0	343	5	BY342594 BY342594
3	33.6	19.0	402	5	BY077721 BY077721
4	33.2	18.8	931	7	CV068875 f2_new_ch
5	33	18.6	348	5	BY338426 BY338426
6	33	18.6	353	5	BY337247 BY337247
7	33	18.6	354	5	BY013584 BY013584
8	33	18.6	355	5	BY340467 BY340467
9	33	18.6	360	5	BY343718 BY343718
10	33	18.6	362	5	BY339252 BY339252
11	33	18.6	371	5	BY083658 BY083658
12	33	18.6	611	7	CF916169 B0991A12-
13	33	18.6	621	5	BQ620008 Talr1138G
14	33	18.6	624	7	CF913548 B095DD02-
15	33	18.6	635	2	AW319335 unilf04.y
16	33	18.6	641	1	AA989838 uc79f07.y
17	33	18.6	657	7	CF915946 B0987F04-
18	33	18.6	669	7	CF174622 B0941F09-
19	33	18.6	973	6	CA490927 AGENCOURT
20	32.8	18.5	638	9	CE066545 tigr-gss-
21	32.6	18.4	786	9	CL421872 ZMMBB043
22	32	18.1	327	6	BY789720 BY789720
23	32	18.1	1297	5	BQ645920 AGENCOURT
24	31.8	18.0	553	7	CR540072 DKFZp459D

C 25 31.8 18.0 1076 5 BX463482
 26 31.6 17.9 1011 9 CL494562
 C 27 31.6 17.9 1310 5 BQ720271
 28 31.4 17.7 344 5 BY340052
 29 31.4 17.7 374 2 BB843476
 C 30 31.4 17.7 681 9 CG295117
 31 31.2 17.6 693 9 CL190317
 32 31.2 17.6 2199 3 CNS0A7F6
 33 31 17.5 829 7 CK192985
 34 30.8 17.4 394 7 CO068301
 C 35 30.8 17.4 446 7 CO320619
 36 30.8 17.4 453 7 CO867977
 C 37 30.8 17.4 639 7 CF886875
 38 30.8 17.4 653 7 CF891128
 39 30.8 17.4 660 6 CB853868
 C 40 30.8 17.4 676 2 BE275414
 41 30.8 17.4 805 8 CC437710
 C 42 30.8 17.4 842 5 BU140827
 43 30.6 17.3 1074 8 CC245979
 44 30.4 17.2 352 5 BY342724
 C 45 30.4 17.2 647 9 AG090615

ALIGNMENTS

RESULT 1
 AQ297217 507 bp DNA linear GSS 15-DEC-1998
 LOCUS HS_2266_A1_A08_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2266 Col=15 Row=A, genomic survey sequence.
 DEFINITION
 ACCESSION AQ297217 GI:4014396
 VERSION AQ297217.1
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 507)
 AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Kellier,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Sequence Tagged Connector
 Plate: 2266 row: A column: 15
 Class: BAC ends
 High quality sequence stop: 507.
 Location/Qualifiers
 1..507
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=2266 Col=15 Row=A"
 /sex="male"
 /note="Organ: sperm; Vector: pBel0BAC11; BAC Clones in E-Coli DH10B"

Query Match 19.1%; Score 33.8; DB 8; Length 507;
 Best Local Similarity 53.4%; Pred. No. 5.1;

Matches 71; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 5 TATCGTGCAGACGCCAGCTTCCTGTTGTGTACCGGAGCGGACGCACTCTTATC 64
 |||||
 Db 106 TTTCCTTTAGTGACACCTACCACTCTGTTGTGCATGAGAATCCGCTATCACCCCTTATG 165
 |||||
 QY 65 GGACAGACGAGCGGCTCCATATACGCGCGCGTTCATCGCGTGTACCGGACGACGAG 124
 |||||
 Db 166 GCACAGACGCTTCCACACGCTGTCGCCAGAGAACATGACCGGCAATTTCAACCAAT 225
 |||||
 QY 125 GCGCCGCGCCGCG 137
 |||||
 Db 226 GTCACTGTCCCTC 238
 |||||

RESULT 2
 BY342594
 LOCUS
 DEFINITION
 BY342594 RIKEN full-length enriched, whole joints Mus musculus cDNA
 clone L230041A03 5', mRNA sequence.

ACCESSION
 BY342594
 VERSION
 BY342594.1 GI:26572082

KEYWORDS
 EST.

SOURCE
 Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 343)

REFERENCE

AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nkaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
 Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
 Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
 Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Kongaya, A.,
 Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
 Ravasi, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
 Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
 Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
 Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shinagawa, A., Yasuniishi, A., Yoshino, M., Waterston, R., Lande, E.S.,
 Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,

Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,

Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,

Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,

Shiraki, T., Tagami, N., Waki, K., Watahiki, A., Muramatsu, M. and

Hayashizaki, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

Tissues were provided by Vassilis Aidinis (Biomedical Sciences
 Research Center 'Al Fleming', Institute of Immunology 14-16 Al
 Fleming street 16672 Vari, Greece) whose assistance we gratefully
 acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.

FEATURES

source

1..343
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clones="L230041A03"
 /tissue_type="whole joints"
 /clone_lib="RIKEN full-length enriched, whole joints"

ORIGIN

Query Match 19.0%; Score 33.6; DB 5; Length 343;
 Best Local Similarity 51.7%; Pred. No. 5.9;
 Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 3 CTTATCTGTGACAGACGCCAGCTTCCTGTTGTGTACCGGAGCGGACGCACTCTTA 62
 |||||
 Db 99 CTTAATGTGATGTCCTCCCTCTTCTTCCTGCTGACGCGGAGACGCGCTTT 158
 |||||
 QY 63 TCAGAACAGACGAGCGGCTCCATATACAGCGCGGCTTATCTCATCGCGGTGACCGGACACG 122
 |||||
 Db 159 CCTCACCAGCTCGACTCTCTCTCGGGGGGAGCAGGCTCGCGTGCCTGCCACCGG 218
 |||||
 QY 123 AGCGCGCGCTCCCGCTTATCGCGCG 147
 |||||
 Db 219 GACAGAGCGTCCCGCTTCTCCCTCTCC 243
 |||||

RESULT 3

BY077721

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BY077721 RIKEN full-length enriched, adult male liver Mus musculus
 cDNA clone K630011C08 5', mRNA sequence.
 BY077721.1 GI:26179135
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 402)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nkaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
 Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
 Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,

Kawaji, H., Kawasawa, Y., Kedzierzki, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravagi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL MEDLINE PUBLISHED
Nature 420, 563-573 (2002)
22354683
12466851

COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Inotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
source
1..402
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="K630011C08"
/sex="male"
/tissue_type="liver"
/dev_stage="adult"
/clone_lib="RIKEN full-length enriched, adult male liver"

ORIGIN
Query Match 19.0%; Score 33.6; DB 5; Length 402;
Best Local Similarity 51.7%; Pred. No. 5.9;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 3 CTTATCGTGACGAGCCAGCTTCCTGTTGCTTAACCGCAGCCGACGACCTCTTA 62
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 89 CTTAATGTGATCCCGCCCCCTTTCTTCCTGAGCTGACGGCGGACGAGCCGCTTT 148

QY 63 TCGGAACAGGAGCGGCTCATATCAGCGCGGCTTATCTCATGCGCGTGACCGGACAG 122
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 149 CTTACCACTCGCACTCTCTCGGGGGCGACGGCTCGCGCTGCCCTGCCACCA 208
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 123 AGCGCGCGCTCGCGCTTATCGGCGC 147

Db 209 GACAGAGCGTCCGCTTCCTCTCC 233
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
CV068875 931 bp mRNA linear EST 24-AUG-2004
LOCUS f2.new chopped.fasta.Contig488 Preamplified custom cDNA library in
DEFINITION pCWSpot6.1 (Resgen, Invitrogen Inc.) Emiliana huxleyi cDNA, mRNA
sequence.
ACCESSION CV068875 GI:51532039
VERSION CV068875.1
KEYWORDS EST.
SOURCE Emiliana huxleyi
ORGANISM Emiliana huxleyi
REFERENCE 1 (bases 1 to 931)
AUTHORS Wahlund, T.W., Zhang, X. and Read, B.A.
TITLE Expressed Sequence Tag Profiles from Calcifying and Non-Calcifying
Cultures of Emiliana huxleyi
JOURNAL Micropaleontology (2004) In press
COMMENT Contact: Betsy Read
Department of Biological Sciences
California State University San Marcos
333 S. Twin Oaks Valley Road, San Marcos, CA 92096-0001, USA
Tel: 760 750 4129
Email: bread@csusm.edu.

FEATURES
source
1..931
/organism="Emiliana huxleyi"
/mol_type="mRNA"
/strain="1516"
/db_xref="taxon:2903"
/dev_stage="Late log growth phase"
/clone_lib="Preamplified custom cDNA library in
pCWSpot6.1 (Resgen, Invitrogen Inc.)"
/note="Emiliana huxleyi grown in Artificial Seawater
(Guillard's F/2 media)."

ORIGIN
Query Match 18.8%; Score 33.2; DB 7; Length 931;
Best Local Similarity 56.4%; Pred. No. 8.2;
Matches 62; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 39 ACCGAGCGGACGCAACTCTTATCGGAACAGGAGCGGCTCCATATCAGCGCGCTT 98
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 425 ACCTGCTCGGCGCGCGCTCTCGCAACAGGCGGAGTCCGCGCAACAGTCCCGC 484
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 99 ATCTCATGCGCTGACCGGACACGAGCGCGCTTCCTTATCGGCGCTT 148
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 485 GGCGCGTGACGAGCGGCGCACTCGGGTCTCGCCGATCGGCTT 534
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
BY338426 348 bp mRNA linear EST 11-DEC-2002
LOCUS BY338426 RIKEN full-length enriched, whole joints Mus musculus cDNA
DEFINITION Clone 12300014G21 5', mRNA sequence.
ACCESSION BY338426 GI:26534820
VERSION BY338426.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 348)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Oato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL

MEDLINE

PUBMED

COMMENT

Nature 420, 563-573 (2002)
12354683
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Vassilis Aidinis (Biomedical Sciences Research Center 'Al. Fleming' Institute of Immunology 14-16 Al Fleming street 16672 Vari, Greece) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

1. 348

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

ORIGIN

Query Match

Best Local Similarity 18.6%; Score 33; DB 5; Length 348;

Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY

3 CTTATCGTGAAGAGCGGCGCTCTCTGTTGTGTTAAACGGCGGAGCGAATCTCTTA 62
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db

99 CTTAATGTGATCCCCCCCCCTTCTCTCGTAGCTGACGGCGGAGCGCGCTTT 158
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY

63 TCGGACAGGAGCGGCGCTCTCATATACGCGCGCGTTATCTCATGCGCGTACGGACAG 122
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db

159 CTTACACGATCGGCACTCTCTCGGGGCGCAGCGCTCGCGTGGCCCTGCCACCGG 218
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY

123 AGGGCGCGCTCCCGCTTATCGCGC 147
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db

219 GACAGAGCTCCGCTTCCCTCTCC 243
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6

BY337247

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

/clone="L230014G21"

/tissue type="whole joints"

/clone_lib=RIKEN full-length enriched, whole joints"

BY337247 353 bp mRNA linear EST 11-DEC-2002
BY337247 RIKEN full-length enriched, whole joints Mus musculus CDNA
clone L230006D15 5', mRNA sequence.

BY337247

BY337247.1 GI:26532790

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (Bases 1 to 353)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Oato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216
 Email: genome-resgsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Vassilis Aidinis (Biomedical Sciences Research Center 'Al. Fleming' Institute of Immunology 14-16 Al. Fleming street 16672 Vari, Greece) whose assistance we gratefully acknowledge.
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

Location/Qualifiers
 1. .353
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="1230006015"
 /tissue_type="whole joints"
 /clone_lib="RIKEN full-length enriched, whole joints"

Query Match 18.6%; Score 33; DB 5; Length 353;
 Best Local Similarity 51.7%; Pred. No. 9.2; Indels 0; Gaps 0;
 Matches 75; Conservative 0; Mismatches 70;

ORIGIN
 3 CTTATCGTCAGCAGCGCCAGCTTCTCTGTTCTTAACCGCAGCGGACGCTCTCTTA 62
 |||||
 99 CTTAATGTGATCCCCCCCCCTTCTCTTCGTTAGCTGACGCGGACGCGCTTT 158
 |||||
 63 TCGGAACAGACGCGGCTCATATACGCGCGGCTTATCTATGCGGTGACGCGACAG 122
 |||||
 159 CCTCACCAGCTCGCACTCTCTCGGGGGCGCAGGCTCGCGCTGCGCCCTGCCACG 218
 |||||
 123 AGCGCGCGCTCGCGTATCGCGCC 147
 |||||
 219 GACAGAGCGTCCGCTTCCCTCTCC 243
 |||||

RESULT 7
 BY013584
 LOCUS
 DEFINITION
 musculus cDNA clone G730040E06 5', mRNA sequence.
 BY013584
 BY013584.1 GI:26073833
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (base 1 to 354)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Baisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Vardaro, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynehaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, B.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 22354683
 12466851

CONTACT: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

Email: genome-resgsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

Location/Qualifiers
 1. .354
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="G730040E06"
 /tissue_type="lung"
 /cell_line="RCB-0558 LLC"
 /clone_lib="RIKEN full-length enriched, lung RCB-0558 LLC cDNA"

FEATURES
 source
 1. .354
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="G730040E06"
 /tissue_type="lung"
 /cell_line="RCB-0558 LLC"
 /clone_lib="RIKEN full-length enriched, lung RCB-0558 LLC cDNA"

ORIGIN

Query Match 18.6%; Score 33; DB 5; Length 354;
 Best Local Similarity 51.7%; Pred. No. 9.2;
 Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 3 CTTATGCTGACAGGAGCGGCTTCTGTGTTGCTAACCGGACGGGACGCAACTCTCTTA 62
 |||||
 Db 116 CTTAATGTGATCCCCCCCCCTTCTCTTCGTTAGCTGACGGGACGCGCGCTTT 175

QY 63 TCGGACAGGAGCGGCTTCCATATCAGCGGCGGCTTATCTATGCGGCTGACCGACAGC 122
 |||||
 Db 176 CTCACACAGCTCGACTTCTTCGGGGGCGCAGCGCTCGCGTGCCTGCCACACG 235

QY 123 AGGCGCCGCTCCGCTTATCGGCGC 147
 |||||
 Db 236 GACAGAGCTCCGCTTCCCTCTCC 260

RESULT 8
 BY340467
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BY340467 355 bp mRNA linear EST 12-DEC-2002
 BY340467 RIKEN full-length enriched, whole joints Mus musculus CDNA
 Clone L230028H13 5', mRNA sequence.
 BY340467
 EST.
 BY340467.1 GI:26569955
 Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 355)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
 Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
 Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
 Gariboldi, M., Giasi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, T.J., Jarvis, E.D., Kanai, A.,
 Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
 Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
 Saitana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
 Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
 Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 22354683
 12466851

CONTACT: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
 Hirozane, F., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
 Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and
 Hayashizaki, Y. Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Vasilis Aidinis (Biomedical Sciences
 Research Center 'Al. Fleming' Institute of Immunology 14-16 Al.
 Fleming street 16672 Vari, Greece) whose assistance we gratefully
 acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.

FEATURES
 Location/Qualifiers
 source
 1. .355
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="L230028H13"
 /tissue_type="whole joints"
 /clone_lib="RIKEN full-length enriched, whole joints"

ORIGIN
 Query Match 18.6%; Score 33; DB 5; Length 355;
 Best Local Similarity 51.7%; Pred. No. 9.2;
 Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 3 CTTATGCTGACAGGAGCGGCTTCTGTGTTGCTAACCGGACGGGACGCAACTCTCTTA 62
 |||||
 Db 99 CTTAATGTGATCCCCCCCCCTTCTCTTCGTTAGCTGACGGGACGCGCGCTTT 158

QY 63 TCGGACAGGAGCGGCTTCCATATCAGCGGCGGCTTATCTATGCGGCTGACCGACAGC 122
 |||||
 Db 159 CTCACACAGCTCGCACTCTCTTCGGGGGCGCAGCGCTCGCGTGCCTGCCACACG 218

QY 123 AGGCGCCGCTCCGCTTATCGGCGC 147
 |||||
 Db 219 GACAGAGCGTCCGCTTCCCTCTCC 243

RESULT 9
 BY343718 360 bp mRNA linear EST 12-DEC-2002
 LOCUS
 DEFINITION
 BY343718 RIKEN full-length enriched, whole joints Mus musculus CDNA
 Clone L230047A12 5', mRNA sequence.
 BY343718
 EST.
 BY343718.1 GI:26573206
 Mus musculus (house mouse)

ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 360)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaado, I., Osato, N., Saito, R., Suzuki, H., Yananaka, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
 Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
 Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,

CONTACT: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
 Hirozane, F., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,

Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reid, J., Reid, J., Ring, B.Z., Ringwald, M., Santelina, A., Schneider, C., Semp, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL
Nature 420, 563-573 (2002)

MEDLINE
22354683

PUBMED
12466851

COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Vassilis Aidinis (Biomedical Sciences Research Center 'Al. Fleming' Institute of Immunology 14-16 Al. Fleming street 16672 Vari, Greece) whose assistance we gratefully acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
FEATURES
source
1. 360
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="1230047A12"
/tissue_type="whole joints"
/clone_lib="RIKEN full-length enriched, whole joints"

ORIGIN
Query Match 18.6%; Score 33; DB 5; Length 360;
Best Local Similarity 51.7%; Pred. No. 9.2;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 3 CTTATCGTACAGGACGCCAGCTTCTGTGTTGCTAAACCGCAGCGGACGCACTCTCTTA 62
|||||
Db CTTAATGATGATCCCGCCCGCTTCTCTTCTCGTAGCTGACGCGGACGACGCGCTTT 158
|||||
QY 63 TCGGACAGGACGCGCTCATATCAGCGCGGCTTATCTATCGCGTGCAGCGGACAGC 122
|||||
Db 159 CCTCACCAGCTCGACTCTCTCGGGGGGCGCAGCGCTCGCGTCCCTGCCCCACCGG 218
|||||
QY 123 AGCGCGCCGCTCCGCTTATCGCGCC 147
|||||
Db 219 GACAGAGCGTCCGCTTCCCTCTCC 243
|||||

RESULT 10
LOCUS
BY339252
DEFINITION
RIKEN full-length enriched, whole joints Mus musculus cDNA
clone 1230021807 5', mRNA sequence.
ACCESSION
BY339252
VERSION
BY339252.1 GI:26568740
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM

REFERENCE
1 (bases 1 to 362)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Oatso, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalia, S., Beisel, K.E., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reid, J., Reid, J., Ring, B.Z., Ringwald, M., Santelina, A., Schneider, C., Semp, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL
Nature 420, 563-573 (2002)

MEDLINE
22354683

PUBMED
12466851

COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Vassilis Aidinis (Biomedical Sciences Research Center "Al. Fleming" Institute of Immunology 14-16 Al. Fleming street 16672 Vari, Greece) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

Location/Qualifiers

1. .362
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="L230021B07"
/tissue_type="whole joints"
/clone_lib="RIKEN full-length enriched, whole joints"

ORIGIN

Query Match 18.6%; Score 33; DB 5; Length 362;
Best Local Similarity 51.7%; Pred. No. 9.2;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 3 CTTATCGTGACAGAGCGGCGTCTCTGTGTGCTAAACGCGGCGGACGCAACTCCCTTA 62
Db CTTAATGTGATCCCCCCCCCTTCTCTTCCTGAGCTGACGCGGCGGCGGCTTT 158

QY 63 TCGGAACAGAGCGCGCTTCATACGCGGCGGCTTATCTATGCGGTGACCGGACG 122
Db CTCACACAGCTGCACTCTCTCGGGGCGGACGCGCTCGCGCTGCGCCGCCACCGG 218

QY 123 AGGCCCGCTCCCGCTTATCGGCGC 147
Db GACAGAGCTCCCGCTTCCCTCTCC 243

RESULT 11

BY083658 LOCUS 371 bp mRNA linear EST 07-DEC-2002
DEFINITION BY083658 RIKEN full-length enriched, 10 days neonate heart Mus musculus cDNA clone K630044L13 5', mRNA sequence.

ACCESSION BY083658

VERSION BY083658.1 GI:26203394

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 371)

AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, B.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,

Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yangisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Arakawa, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hata, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Wataniki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

Location/Qualifiers

1. .371

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="K630044L13"

/tissue_type="heart"

/dev_stage="10 days neonate"

/clone_lib="RIKEN full-length enriched, 10 days neonate heart"

ORIGIN

Query Match 18.6%; Score 33; DB 5; Length 371;
Best Local Similarity 51.7%; Pred. No. 9.2;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 3 CTTATCGTGACAGAGCGGCGTCTCTGTGTGCTAAACGCGGCGGACGCAACTCCCTTA 62
Db CTTAATGTGATCCCCCCCCCTTCTCTTCCTGAGCTGACGCGGCGGCGGCTTT 170

[illegible]

ORIGIN

constructed by Yulan Piao."

Query Match 18.6%; Score 33; DB 7; Length 624;
 Best Local Similarity 51.7%; Pred. No. 9.3;
 Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 3 CTTATCGTGACAGGACGCGAGCTTCCTGTGTTCTTAACGCGACGCGACGCACTCCTTA 62
 Db CTTAAATGTGATCCCCCCCCCTTTTCCTTTCTCGTAGCTTGACGCGGACGCGCTTT 92
 QY 63 TCGGAACAGGACGCGCTCCATATCAGCGCGGTTATCTCATGCGGTGACCGGACAG 122
 Db 93 CTTACACAGCTCGCATCTCTCTCGGGGGGCGACGCGTCCGCTGCTCCACACGG 152
 QY 123 AGCGCGCGTCCCGCTTATCGCGCC 147
 Db 153 GACAGAGCGTCCCGCTTCCTCTCC 177

RESULT 15
 AW319335
 LOCUS
 DEFINITION
 unilf04.y1 Sugano mouse kidney mkoa Mus musculus cDNA clone
 IMAGE:2395039 5' similar to TR:Q9YGP5 Q9YGP5 RRM-TYPE RNA-BINDING
 PROTEIN HERMES. ;, mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AW319335
 AW319335.1 GI:6748879
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 635)
 Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
 Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
 Person,B., Waller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
 Ritter,B., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
 Waterston,R. and Wilson,R.
 The WashU-NCI Mouse EST Project 1999
 Unpublished (1999)
 Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:1013187

TITLE
 JOURNAL
 COMMENT

Seq primer: custom primer used
 High quality sequence stop: 305.
 Location/Qualifiers
 1. 635
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:2395039"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse kidney mkoa"
 /note="Organ: Kidney; Vector: pME18S-FL3; Site 1: DralIII
 (CACTGTGTG); Site 2: DraIII (CACTATGTG); 1st strand cDNA
 was primed with an oligo(dT) primer
 [ATGTGGCGCTTTTCTTTTCTTTT]; double-stranded cDNA was
 ligated to a DraIII adaptor [TCTTGGCTACTGG], digested
 and cloned into distinct DraIII sites of the pME18S-FL3
 vector (5' site CACTGTGTG, 3' site CACTATGTG). XhoI should
 be used to isolate the cDNA insert. Size selection was
 performed to exclude fragments <1.5kb. Library
 constructed by Dr. Sumio Sugano (University of Tokyo

FEATURES
 source

ORIGIN

Search completed: May 10, 2005, 07:25:12
Job time : 1244.12 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 03:40:34 ; Search time 3416.8 Seconds
(without alignments)
7998.346 Million cell updates/sec

Title: US-09-896-888A-1
Perfect score: 564
Sequence: 1 catgatgataacaatgtat.....tggtagcgacacaacatg 564

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	564	100.0	564	6	BD070856 Insect ex
2	564	100.0	4170	14	M83827 Orgyia pseu
3	564	100.0	131995	14	U75930 Orgyia pseu
C 4	562.4	99.7	1429	14	S64501 p8.9-8.9 kd
5	548	97.2	2773	6	AX766573 Sequence
C 6	111.8	19.8	118584	14	AY043265 Epiphyas
C 7	101.8	18.0	131526	14	AY145471 Rachiplus
C 8	101.4	18.0	1511	14	M62488 Autographa
C 9	101.4	18.0	133894	6	A48542 Sequence 1
C 10	101.4	18.0	133894	14	L22858 Autographa
C 11	100	17.7	131158	14	AY327402 Choriston
C 12	93.4	16.6	28413	6	BD187790 A virus i
C 13	93.4	16.6	128413	14	NPH3COMP
C 14	92.2	16.3	129609	14	AF512031 Choriston
C 15	92	16.3	2178	14	NPB8MIEN
C 16	90.4	16.0	2011	14	M59422 Autographa
17	73.8	13.1	155060	14	MBU59461
18	72.2	12.8	4051	14	AF246708 Spodopter
C 19	72.2	12.8	139342	14	AF325155 Spodopter

20	72.2	12.8	153656	14	AF539999	Mamestra
C 21	69.2	12.3	15528	14	AF107100	AF107100 Ecotropis
22	68.2	12.1	158482	14	AY126275	AY126275 Mamestra
C 23	67.4	12.0	130759	14	AF303045	AF303045 Helicoverpa
C 24	66.4	11.8	14235	14	HZU67264	U67264 Helicoverpa
C 25	65.6	11.6	130869	14	AF334030	AF334030 Helicoverpa
C 26	65.4	11.6	135611	14	AF169823	AF169823 Spodopter
C 27	64	11.3	131403	14	AF271059	AF271059 Helicoverpa
C 28	63.6	11.3	161046	14	AF081810	AF081810 Lymantria
C 29	59.2	10.5	113220	14	AP066270	AP066270 Adoxophye
C 30	41.8	7.4	131158	14	AY327402	AY327402 Choriston
C 31	40.2	7.1	128380	2	AC151093	AC151093 Bos taurus
C 32	39.6	7.0	6289	6	AX598859	AX598859 Sequence
C 33	39.6	7.0	9289	6	AX251255	AX251255 Sequence
C 34	39.6	7.0	9289	6	AX767469	AX767469 Sequence
C 35	39	6.9	152573	2	CR376837	CR376837 Danio rer
C 36	39	6.9	200989	5	BX000999	BX000999 Zebrafish
C 37	38	6.7	6289	6	AX599005	AX599005 Sequence
C 38	38	6.7	9289	6	AX767545	AX767545 Sequence
C 39	38	6.7	160376	4	AC150919	AC150919 Bos taurus
C 40	37.4	6.6	96824	2	AC020384	AC020384 Drosophil
C 41	36.6	6.5	125422	9	AL449063	AL449063 Human DNA
42	36.6	6.5	175909	9	AC093559	AC093559 Homo sapi
C 43	36.6	6.5	190808	10	AL928537	AL928537 Mouse DNA
C 44	36.6	6.5	221128	10	AC123846	AC123846 Mus muscu
C 45	36.4	6.5	505	6	AR424142	AR424142 Sequence

ALIGNMENTS

RESULT 1	BD070856	564 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD070856				
DEFINITION	Insect expression vectors.				
ACCESSION	BD070856				
VERSION	BD070856.1 GI:22616459				
KEYWORDS	JP 2001516225-A/1.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 564)				
AUTHORS	Grigliatti,T.A., Theilmann,D.A., Pfeifer,T.A. and Hegedus,D.D.				
TITLE	Insect expression vectors				
JOURNAL	Patent: JP 2001516225-A 1 25-SEP-2001;				
COMMENT	THE UNIVERSITY OF BRITISH COLUMBIA				
	OS Multicapsid nucleopolyhedrovirus				
	PN JP 2001516225-A/1				
	PD 25-SEP-2001				
	PF 26-MAR-1998 JP 1998541010				
	PR 27-MAR-1997 US 60/049946,28-JAN-1998 CA 2221819 PI				
	THOMAS A GRIGLIATTI,DAVE A THEILMANN,THOMAS				
	A PFEIFER,DWAYNE D				
	PI HEGEDUS				
	PC C12N15/85,C12N5/06,C12N15/69//C12N9/22				
	CC Insect expression vectors				
	FT Key Location/Qualifiers				
	FT source 1..564				
	FT Location/Qualifiers				
	1..564				
	/organism="unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				

FEATURES

source

ORIGIN

Query Match 100.0%; Score 564; DB 6; Length 564;
Best Local Similarity 100.0%; Pred. No. 4e-171;
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 CATGATGATAACAATGTATGTGCTTCAACAACAATTCGTGAACTGTG 60
|||||
1 CATGATGATAACAATGTATGTGCTTCAACAACAATTCGTGAACTGTG 60

DB

```
QY 61 TTTTCATGCTTTGCCAACAAGCACCTTTATATCTCGGTGGCCTCCCAACCAACTTTTTT 120
Db 61 TTTTCATGCTTTGCCAACAAGCACCTTTATATCTCGGTGGCCTCCCAACCAACTTTTTT 120
QY 121 GCATGCAAAAACACACGCTTTTGACGCGGGCCCATACATAGTACAACTCTAGTTTC 180
Db 121 GCATGCAAAAACACACGCTTTTGACGCGGGCCCATACATAGTACAACTCTAGTTTC 180
QY 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGTTCCAAATACACTACCAC 240
Db 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGTTCCAAATACACTACCAC 240
QY 241 ATTGAACCTTTTTCAGTGCAAAAAGTACGTGCGGAGTACGTAGCGCGGCTTATC 300
Db 241 ATTGAACCTTTTTCAGTGCAAAAAGTACGTGCGGAGTACGTAGCGCGGCTTATC 300
QY 301 GGGTCGCTCTCTGTCCAGTACGAATACATATCGGACCGGACGAGTGTCTTATCGT 360
Db 301 GGGTCGCTCTCTGTCCAGTACGAATACATATCGGACCGGACGAGTGTCTTATCGT 360
QY 361 GACAGGACGCGCTCTCTGTCTTAACCGGACCGGACGCAACTCTTATCGGAACA 420
Db 361 GACAGGACGCGCTCTCTGTCTTAACCGGACCGGACGCAACTCTTATCGGAACA 420
QY 421 GGACGCGCTCCATATACGCGCGGTATCTCATGCGCGTACCGGACGAGCGGCC 480
Db 421 GGACGCGCTCCATATACGCGCGGTATCTCATGCGCGTACCGGACGAGCGGCC 480
QY 481 GTCCGCTTATCGCGCTTAAATACAGCCCGGACGATCTGGTAAACACAGTTGAACAG 540
Db 481 GTCCGCTTATCGCGCTTAAATACAGCCCGGACGATCTGGTAAACACAGTTGAACAG 540
QY 541 CATCTGTTACGCGACACACATG 564
Db 541 CATCTGTTACGCGACACACATG 564

RESULT 2
NHPTAA 4170 bp DNA linear VRL 02-AUG-1993
LOCUS Orgyia pseudotsugata nuclear polyhedrosis virus transcriptional
DEFINITION trans-activator (IE-2) gene, complete cds; ORF, complete cds.
ACCESSION M83827
VERSION M83827.1 GI:332540
KEYWORDS transcriptional transactivator.
SOURCE Orgyia pseudotsugata single capsid nucleopolyhedrovirus
ORGANISM Orgyia pseudotsugata single capsid nucleopolyhedrovirus
          Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
          Nucleopolyhedrovirus.
REFERENCE 1 (bases 1 to 4170)
AUTHORS Theilmann,D.A. and Stewart,S.
TITLE Molecular analysis of the trans-activating IE-2 gene of Orgyia
JOURNAL pseudotsugata multicapsid nuclear polyhedrosis virus
Virology 187 (1), 84-96 (1992)
MEDLINE 92142536
PUBMED 1736546
REFERENCE 2 (bases 1 to 4170)
AUTHORS Theilmann,D.A. and Stewart,S.
TITLE Tandemly repeated sequence at the 3' end of the IE-2 gene of the
JOURNAL baculovirus Orgyia pseudotsugata multicapsid nuclear polyhedrosis
MEDLINE 92142537
PUBMED 1736547
COMMENT Original source text: Orgyia pseudotsugata nuclear polyhedrosis
virus DNA.
FEATURES
    source
        1. 4170
            /organism="Orgyia pseudotsugata single capsid
            nucleopolyhedrovirus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10450"
```

gene

```
TATA_signal
mRNA
CDS
627..3959
/gene="IE-2"
627..634
/gene="IE-2"
657..21908
/gene="IE-2"
691..1908
/gene="IE-2"
/codon_start=1
/product="transcriptional trans-activator"
/protein_id="AAA46749.1"
/db_xref="GI:332541"
```

```
/translation="MSRSNNAPTPSNRRRLSLVRGRLLTSPPPDAASAQRASPPR
SAPRAAPRVHAGVDCGAPLARASVALPVGYNLHGDHFNPPREDDDDILFVDTAAEOA
RQANVLHEAVNRHERLRRLGERMTRSPILLNYSYSPSTPSRSRSPDLIMPEDLQP
ARQPLAPFNDDDDERLLEQVLESAAFPQAPPAQVDSVLCISCIFDILQ
NYSNFSVTSTECNHAFCYKVSIVFGKYSICNRRTTISCRAYNRDGVLELSTMS
TVNDSQAIKRWALQSDSNPHSNMTTIOELQAEALRAAATARAHDVNMARSDSQ
LLRQQLDVKEALAHESNARLKLQKQNETLSANNLSLQHLNTQVIESRVMEQFKRQ
HDEMEKFKLSLS"
```

polyA_signal

```
1958..1963
/gene="IE-2"
2017..2759
/gene="IE-2"
2915..2922
/gene="IE-2"
2944..3907
/gene="IE-2"
2984..3907
/gene="IE-2"
```

```
/note="ORF, homologous to the ACMNPV PE-38 gene; putative"
```

```
/codon_start=1
```

```
/protein_id="AAA46750.1"
```

```
/db_xref="GI:332542"
```

```
/translation="MSRYSPIRNVGRRRRVQERLLSBNFNSRPVTAVAACVACLE
TYCQSNIIIDFLMPSCTHLFCYKVLNNYKNAMNVPRAAVSCPMCKKVGTVQAF
PNSVWSCFKTKTGDRTPACLFQMLAKTIQDRYMAEEAEETPSFVINKLQALDQ
AOKEARDLQESMERQKOHNVAVNSCEQVTLQTLADMOALDRSEALSSTLAEHN
RAANVOIDLSRRRAVORLEAAQAPVSVNVEFNQNTNLHERFSRVSVTVSDMMI
EDSKSLQSHVFGAACLPCSVNVEINFPPE"
```

```
3954..3959
/gene="IE-2"
```

polyA_signal

ORIGIN

```
Query Match 100.0%; Score 564; DB 14; Length 4170;
Best Local Similarity 100.0%; Pred. No. 6.1e-171;
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGATGATAAACAAATGATGCTGCTAATGTTGCTTCAACAACTTCTGTTGAACGTG 60
Db 130 CATGATGATAAACAAATGATGCTGCTAATGTTGCTTCAACAACTTCTGTTGAACGTG 189
QY 61 TTTTCATGTTTGCACAAAGCAGCTTTTATCTCGGTGGCCTCCCAACCACTTTTTT 120
Db 190 TTTTCATGTTTGCACAAAGCAGCTTTTATCTCGGTGGCCTCCCAACCACTTTTTT 249
QY 121 GCATGCAAAAACACGCTTTTGACGCGGGCCCATACATAGTACAACTCTAGTTTC 180
Db 250 GCATGCAAAAACACGCTTTTGACGCGGGCCCATACATAGTACAACTCTAGTTTC 309
QY 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGTTCCAAATACACTACCAC 240
Db 310 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGTTCCAAATACACTACCAC 369
QY 241 ATTGAACCTTTTTCAGTGCAAAAAGTACGTGCGGAGTACGTAGCGCGGCTTATC 300
Db 370 ATTGAACCTTTTTCAGTGCAAAAAGTACGTGCGGAGTACGTAGCGCGGCTTATC 429
QY 301 GGGTCGCTCTCTGTCCAGTACGAATACATATCGGACCGGACGAGTGTCTTATCGT 360
Db 430 GGGTCGCTCTCTGTCCAGTACGAATACATATCGGACCGGACGAGTGTCTTATCGT 489
QY 361 GACAGGACGCGCTTCTCTGTGTTGTTAAACCGGACCGGACGCAACTCTTATCGGAACA 420
```

```
Db 490 GACAGACGCCAGCTTCTGTGTGTTGTAACCGCAGCGGACCACTCTTATCGGAACA 549
|
Qy 421 GGACGGCGCTCCATATCAGCGCGCGCTTATCTATCGCGGTGACCGACACAGGCGGCC 480
|
Db 550 GGACGGCGCTCCATATCAGCGCGCGCTTATCTATCGCGGTGACCGACACAGGCGGCC 609
|
Qy 481 GTCCCGCTTATCGCGCTTAAATACAGCGCGCAACGATCTGTGTAACACAGTTGAACAG 540
|
Db 610 GTCCCGCTTATCGCGCTTAAATACAGCGCGCAACGATCTGTGTAACACAGTTGAACAG 669
|
Qy 541 CATCTGTTACAGCGCACACATG 564
|
Db 670 CATCTGTTACAGCGCACACATG 693
|

RESULT 3
OPU75930 131995 bp DNA circular VRL 25-MAR-2003
LOCUS Orygia pseudotsugata multicapsid nucleopolyhedrovirus, complete
DEFINITION
ACCESSION U75930.2 GI:11024985
VERSION
KEYWORDS
SOURCE Orygia pseudotsugata multicapsid nucleopolyhedrovirus
ORGANISM Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
Nucleopolyhedrovirus.
REFERENCE 1 (bases 1 to 131995)
AUTHORS Ahrens,C.H., Russell,R.L., Funk,C.J., Evans,J.T., Harwood,S.H. and
Rohrmann,G.F.
TITLE The sequence of the Orygia pseudotsugata multinucleocapsid nuclear
JOURNAL polyhedrosis virus genome
MEDLINE Virology 229 (2), 381-399 (1997)
PUBMED 9126251
REFERENCE 2 (bases 1 to 131995)
AUTHORS Rohrmann,G.F.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1996) Oregon State University, Agricultural
Chemistry, Corvallis, OR 97331-7301, USA
REFERENCE 3 (bases 1 to 131995)
AUTHORS Rohrmann,G.F.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-1998) Oregon State University, Agricultural
Chemistry, Corvallis, OR 97331-7301, USA
REMARK Sequence update by:submitter
AUTHORS 4 (bases 1 to 131995)
TITLE Direct Submission
JOURNAL Submitted (26-OCT-2000) Oregon State University, Agricultural
Chemistry, Corvallis, OR 97331-7301, USA
REMARK Sequence update by submitter
COMMENT On Oct 26, 2000 this sequence version replaced gi:2934903.
FEATURES
source
1. 131995
/organism="Orygia pseudotsugata multicapsid
nucleopolyhedrovirus"
/mol_type="genomic DNA"
/db_xref="taxon:164623"
/notes="similar to Autographa californica nuclear
polyhedrosis virus (AcMNPV) deposited in GenBank Accession
Number L22858"
complement(123..947)
/notes="ORF1; pk-1; similar to AcMNPV ORF10"
/codon_start=1
/product="protein kinase 1"
/protein_id="AAC59000.1"
/db_xref="GI:1911247"

CDS
/translation="MDAALQSLSRFVADCAVLAPKVNGRFGQMDVLRHPTTSKFL
RTKTAHSFSAIDINVDHLSDFSVDMYFCYSSPTAWAIVMDYVPCDPLFELTQ
GALDNLAVNIVRLQDALNDLHNTGYIHNDVKLENLYFGARDRVLYCDVGLCKRE
HSPVHDTLEYFPEKIRRHNYARSFDWAVGLAYKLITGGRHFPERSADEVIDLAS
```

```
CDS
/translation="MMRQYNDPAAALKNVRLNMARDVFVFLTRFNFECSRSTDYKQIAKHSFLASRDYI"
946..2370
/note="ORF2; similar to AcMNPV ORF9"
/codon_start=1
/product="1629-capsid"
/protein_id="AAC59001.1"
/db_xref="GI:2934904"
/translation="MMRQYQSVRSYLYINNOHNAIAAGPFLQRVAGPEAHSVGRNVGD
RAVLRNQAVLDLLKLAEDYADTAYMQADQPEASRSHFATLNRMLLIGVQDDPAR
RNLNSVLARI EALRLVVDVAENVLGDYEEYSKYISYQITFAQTPTASQOQTQ
TSLRPQTSLSRQTSASLPQPTFDQPEMVSPSPFVHTTPTAILPQTQPTATDFSRP
SDFVYVPGKERAVDTRFKPVPKPEHLKSRSPSVATNAAGATPVA PPRPPPSAD
VTSMPPPPPPPSADVTSMPPPPPMVDLATSMPPPPPPPMVDLATSMPPPINNA
INLLIDAMVAETNKNAGNSALLDQIKGTLTKTKTQADGAPATDPSTLSELRQ
GKTLKLRKIEDQSTQTLTKDQVDTTKTKTLTKNFVTNIDIRISKEQOEKORLDTIT
KRPVAVHTDGNSTGNSDDMRDD"
complement(2533..3270)
/notes="ORF3; ph; similar to AcMNPV ORF8"
/codon_start=1
/product="polyhedrin"
/protein_id="AAC59002.1"
/db_xref="GI:1911249"
/translation="MPDYSYRPTIGRTYVVDNKKYKNGLSVKNKAKRKKHLEHDE
KHLDFDHYVAEDFELGEGKQKLTFLKEIRNVKPDTKLIVNWSGKEFLERTWTRF
VEDSPFVNDQEVMDVFLVNMRTPRNRCYKFLAQHARLWDCDYVPHEVIRIVEPSY
EILMEYRISLAKGGGCGCPINNIHAYTNSFESFVNRVWENFYKPIVYIGTDSSEE
EILMEYSLVFKVEFAPDAPLFTGPAY"
complement(3350..3541)
/notes="ORF4"
/codon_start=1
/product="unknown"
/protein_id="AAC59003.1"
/db_xref="GI:1911250"
/translation="MRKNGASQYPEHLTTVHKRHLQLSCVWCDDQSRWDPHSAKGL
MFEHMFCLRIYVQKIKH"
complement(3698..4093)
/notes="ORF5"
/codon_start=1
/product="unknown"
/protein_id="AAC59004.1"
/db_xref="GI:1911251"
/translation="MGASQNSQKILKLFYRWSSQTGAALDDEKDLHLYDLERFVGA
HLNKRADDKARKKCAKRAALKRVEIADRHLMEAAAPACADDGSRWSTLSRAQLDDI
AREKEIVDIRHLRLQLKQDSLKKQDLKKR"
complement(4125..4739)
/notes="ORF6; left-2; similar to AcMNPV ORF6"
/codon_start=1
/product="late expression factor 2"
/protein_id="AAC59005.1"
/db_xref="GI:1911252"
/translation="MERYVNPAGIDGLKRSYLYVDPHDFVGLTSLPVTVPERGLF
VMSGMRLLALLAAAPQEPQPAVRFPQRSRRNVCCLKACAGAQSLAKVLAARVMP
PCSKMTADLSSAPNRNKRKPFENYLANVITCTKCTACGLIALLHFRMDAKCV
GSEVTHLLIAQDVYKPSNCAKMKVTKLCPQASCKGLNPICNF"
complement(4742..4975)
/notes="ORF7; similar to AcMNPV ORF5"
/codon_start=1
/product="unknown"
/protein_id="AAC59006.1"
/db_xref="GI:1911253"
/translation="MNRPTMRNTAAVTTDYDREQLRRLNLSRRSVHELCTRSATGFD
CNRFLEAGDRAPAVIVKAAANGQHSSLLCDKV"
5020..5460
/notes="ORF8; similar to AcMNPV ORF4"
/codon_start=1
/product="unknown"
/protein_id="AAC59007.1"
/db_xref="GI:1911254"
/translation="MPGARFVFSRLTQEPKENVVAHVHLMGLRALIDGKVTSDV
RRFGLSNRVLSACVANNVQYATDPTIDMRHQPTIYFRVQCNCHAMADVAPDDHS
IARYLLAECGAVLVDHPDLDFVFGETEGBEGVNELLEVRINAGGDL"
complement(5669..6151)
/notes="ORF9; putp-2"
```

```
/codon_start=1
/product="protein tyrosine phosphatase 2"
/protein_id="AAC59008.1"
/db_xref="GI:1911255"
/translation="MYEASQDEHVFVGGYGDNEAMLRFIENHAIASVISLIDSDVA
PIRALGHPVGDHIVHYCEAAPCTCAALFNAMPALYDYMVRRIETGEGKRVLHVCYAGASR
SAALAVYLMRSRQMPYQDALNAVQSKRVAINDHFVRLATRSYRFPVNDLKPQIV
"
complement(6129..6791)
/note="ORF10; ptp-1; similar to AcMNPV ORF1"
/codon_start=1
/product="protein tyrosine phosphatase 1"
/protein_id="AAC59009.1"
/db_xref="GI:1911256"
/translation="MLTSRRWAVIYTRISERNCGAWCTNGVRRRRQVHLSPVRAM
FDRWHEYTAGAVTEGRLLJCFKVPFLNAELFEYVTSDEDRWTAASVLARHSALGAVI
DLTNTARYDGAQMVKMLLYKKIRVPGRAVPDDDI VAEFTETVDEFFRRCPPTMLVAV
HWTHGLNRSGLVLCRYMVERLGVSPTDAIARFETARGHKIERTNYLQDLLARKHVRGQ
PN"
6819..7814
/note="ORF11; similar to AcMNPV ORF11"
/codon_start=1
/product="unknown"
/protein_id="AAC59010.1"
/db_xref="GI:1911257"
/translation="WLSLSKLLVAVYGSYNLPHDRYGESVHLVRIVHEHILTNTVVS
ASCVRDIATACLSNHLCPDVARQLLDVSEVAARLSAMFRGSDATGLCADMORALA
DIDHAPLARRVGRANIFALDAADIPTSDVNNLQGLIGRPFHFRPGSLGARVADV
DPDIRADGWYHKFCVLTVMHLVACGAPGASATRLRDVAKHGIPNDGNCAPAIAA
VYGRFCATGREHFAHAKMHILFQFMRNDLTPADERHPCFGVHKDFGRCKDTYTD
LRTHADALIHGTDRQNALFDLLCCVNASDIDADCYCVNVNFYATQNKYKM"
complement(8180..9142)
/note="ORF12; similar to AcMNPV ORF13"
/codon_start=1
/product="unknown"
/protein_id="AAC59011.1"
/db_xref="GI:1911258"
/translation="WFSWFGWNAADQVNAEDQAYREYAVDQRAHSLDVRDVE
RCHPFVFRVYLDTEAGRCSSVDVFCGLKLSHDLQRCNDFRQHVQLNELVLGAP
PAQPSGLSATRHGLVQLLOQVDFKANKNEVLAVGANKDRDNDLDDKLEAVLNHV
KTLNNSDKFISAKHSFKEVGARFEQERQLDITDKNALQCAAPRTAPGVVFFR
DVTKPHLAVFRVGVDTQIAFARGQEEHFRKLEFEEGMDVDVGRAPNPLLA
VHCIEEPANGSHKIRLPPKKVIEVDCAVNVAKDIVKAILNKT"
complement(9094..9825)
/note="ORF13; left-1; similar to AcMNPV ORF14"
/codon_start=1
/product="late expression factor 1"
/protein_id="AAC59012.1"
/db_xref="GI:1911259"
/translation="WAPCKYTPERVOMMDAIVNDSRLAFMTDRPWWAHNFEDS
AAQLFALYVKNISDVHYKPLEEGREWVIDADFKDCADKAEMLKYNVGTATMFLFF
EGKEDAVORINFSNGRPHLMKFCGKFMADPKSLREHFNWFKPAKLVSIGDIRPG
SFADCVRRAVHYIGDAREDLVLRYPDVRDVFNCNANKQIRAPFSYNYKGGYSRCL
TQQLQQRITACASAGSLAGGTPTTSK"
9893..11362
/note="ORF14; egt; similar to AcMNPV ORF15"
/codon_start=1
/product="ecdysteroid UDP-glucosyltransferase"
/protein_id="AAC59013.1"
/db_xref="GI:1911260"
/translation="WVFLIALTLATGARAASILAVPTPAYSHVVYVYAVHALVK
NADNVTVIKPOLLVAVQDEGRVEQIDADMSAQYKKLVASGVRKRGVVADETVI
TCHNMYKLEMPKQFDNANVRFLSTNRTFDAVVEAPADYALVFGHLFPAPVIOI
APGYLAENFERRARAPLHPYPTFGAAALTRGGALSEWRLNEFELLARRSDELL
QKQFGKSTPTTQRURDNVQLLLNLHPVDNNRPVPSVQYLGGGLHQAQLPORDA
PLERLNDSGVAVYSGSDITNSIHAEFLQMLDTPFANLNNYTLVLMKYDDAAS
VALPRNLVAQWFSOTVLRKHNVAVFVTOAGLOSDEALQARVPMVCLPMWQGFH
ARKLQFCVARALDTAAVAPQLOLAIREVIADGEAYEARIDKLRAVHEHDAAPDEKA
VKFTERVTKFNNVWNPARSLLKTTAANWAYSDYVRFPL"
11547..12140
/note="ORF15; similar to AcMNPV ORF16"
/codon_start=1
```

```
Query Match      100.0%; Score 564; DB 14; Length 131995;
Best Local Similarity 100.0%; Pred. No. 1.3e-170;
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGATGATAAACAAATGTATGGTCTAATGTGTCTTCAACAACAATCTCTGTGAACCTGTG 60
Db 128172 CATGATGATAAACAAATGTATGGTCTAATGTGTCTTCAACAACAATCTCTGTGAACCTGTG 128231

QY 61 TTTTCATGTTTGGCAACAAAGCACCTTTATCTCGTGGCGCTCCCAACACCAACTTTTTT 120
Db 128232 TTTTCATGTTTGGCAACAAAGCACCTTTATCTCGTGGCGCTCCCAACACCAACTTTTTT 128291

QY 121 GCACATGCAAAAAACACAGCTTTTGCAGCGCGGCCCATACATAGTACAAACTCTACGTTTC 180
Db 128292 GCATGCAAAAAACACAGCTTTTGCAGCGCGGCCCATACATAGTACAAACTCTACGTTTC 128351

QY 181 GTAGACTATTTTACATAAATAGTCTACACCGCTTGTATACGCTCCAAATACACTACCACAC 240
Db 128352 GTAGACTATTTTACATAAATAGTCTACACCGCTTGTATACGCTCCAAATACACTACCACAC 128411

QY 241 ATTGAACCTTTTTCAGTGCAGGCAAAAAGTACGTGTGGCAGTCACGTAGGCGGCGCTTATC 300
Db 128412 ATTGAACCTTTTTCAGTGCAGGCAAAAAGTACGTGTGGCAGTCACGTAGGCGGCGCTTATC 128471

QY 301 GGGTCGGCTCTGTTCACGTACGAATCACATTATCGACCGGACGAGTGTGTCTTATCGT 360
Db 128472 GGGTCGGCTCTGTTCACGTACGAATCACATTATCGACCGGACGAGTGTGTCTTATCGT 128531

QY 361 GACAGGACGCGCAGCTTCTGTGTGCTAACCGCAGCGGACGCAACTCTCTTATCGGAACA 420
Db 128532 GACAGGACGCGCAGCTTCTGTGTGCTAACCGCAGCGGACGCAACTCTCTTATCGGAACA 128591

QY 421 GGACGGCGCTCCATATCAGCGCGCGGTATCTCATGCGGTGACCGGACGAGCGCGCC 480
Db 128592 GGACGGCGCTCCATATCAGCGCGCGGTATCTCATGCGGTGACCGGACGAGCGCGCC 128651

QY 481 GTCCCGCTTATCGCGCTTATAAATACAGCCGCAACGATCTGTGTAACACAGAGTTGAACAG 540
Db 128652 GTCCCGCTTATCGCGCTTATAAATACAGCCGCAACGATCTGTGTAACACAGAGTTGAACAG 128711

QY 541 CATCTGTTTACAGCGACCAACATG 564
Db 128712 CATCTGTTTACAGCGACCAACATG 128735

RESULT 4
S64501/c
LOCUS
DEFINITION      S64501      1429 bp      DNA      linear      VRL 30-SEP-1993
                  nuclear polyhedrosis virus OpMNPV, Genomic, 1429 nt.
ACCESSION      S64501
VERSION        S64501.1  GI:404518
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS      Wu,X., Stewart,S. and Theilmann,D.A.
TITLE      Characterization of an early gene coding for a highly basic 8.9K
                  protein from the Orgyia pseudotsugata multicapsid nuclear
                  polyhedrosis virus
JOURNAL      J. Gen. Virol. 74 (Pt 8), 1591-1598 (1993)
MEDLINE      93346965
PUBMED      8345350
REMARK      GenBank staff at the National Library of Medicine created this
                  entry [NCBI gibbsq 136152] from the original journal article.
                  Location/Qualifiers
FEATURES
source
organism="Orgyia pseudotsugata multicapsid
nucleopolyhedrovirus"
/mol_type="genomic DNA"
/db_xref="taxon:164623"
```

gene	763. .990	(DK) ; Bratt, Tomas (DK) ; Voldborg, Bjorn (DK) ; Mouritsen, Soren (DK)			
CDS	/gene="p8.9" /note="8.9 kda basic protein" 763. .990 /gene="p8.9" /note="8.9 kda basic protein" /codon_start=1 /protein_id="AAB27738.1" /db_xref="GI:404519" /translation="NNSWQIRMAKQOQVRVAROHRRAAKLGRLYKAKKRLAECEKLG LQRVNDAALAKAPEEFVYPHFSPLYTLN"	Location/Qualifiers 1. .2773 /organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630" /note="p2ZOp2f expression vector for insect cells" 561. .566 /note="HindIII site" 573. .578 /note="Aval site" 586. .591 /note="EcoRI site" 593. .598 /note="BamHI site" 625. .630 /note="ClaI site" 629. .634 /note="ClaI site" 1156. .1161 /note="ApaLI site" 2128. .2133 /note="PstI site" 2204. .2209 /note="NcoI site" 2284. .2289 /note="Aval site" 2294. .2299 /note="Aval, SmaI, and XmaI site" 2551. .2556 /note="ApaLI site"			
ORIGIN	Query Match Best Local Similarity 99.7%; Score 562.4; DB 14; Length 1429; Matches 563; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	Query Match Best Local Similarity 97.2%; Score 548; DB 6; Length 2773; Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 CATGATGATAAACAATGATGCTAATGCTTCAACAACAATTCGTGTAAGTGTG 60	QY	1 CATGATGATAAACAATGATGCTAATGCTTCAACAACAATTCGTGTAAGTGTG 60		
Db	765 CATGATGATAAACAATGATGCTAATGCTTCAACAACAATTCGTGTAAGTGTG 706	Db	5 CATGATGATAAACAATGATGCTAATGCTTCAACAACAATTCGTGTAAGTGTG 64		
QY	61 TTTTCATGTTTGCACAAGCAGCTTTATCTACGCGGGCCCTCCCAACCACTTTTTC 120	QY	61 TTTTCATGTTTGCACAAGCAGCTTTATCTACGCGGGCCCTCCCAACCACTTTTTC 120		
Db	705 TTTTCATGTTTGCACAAGCAGCTTTATCTACGCGGGCCCTCCCAACCACTTTTTC 646	Db	65 TTTTCATGTTTGCACAAGCAGCTTTATCTACGCGGGCCCTCCCAACCACTTTTTC 124		
QY	121 GCACGCAAAAAAACAACGCTTTTGACGCGGGCCCATACATAGTACAAACTCTACGTTTC 180	QY	121 GCACGCAAAAAAACAACGCTTTTGACGCGGGCCCATACATAGTACAAACTCTACGTTTC 180		
Db	645 GCACGCAAAAAAACAACGCTTTTGACGCGGGCCCATACATAGTACAAACTCTACGTTTC 586	Db	125 GCACGCAAAAAAACAACGCTTTTGACGCGGGCCCATACATAGTACAAACTCTACGTTTC 184		
QY	181 GTAGACTATTTTACATAAATAGTCTACACGCTTGATAGCTTCCAAATACACTACACAC 240	QY	181 GTAGACTATTTTACATAAATAGTCTACACGCTTGATAGCTTCCAAATACACTACACAC 240		
Db	585 GTAGACTATTTTACATAAATAGTCTACACGCTTGATAGCTTCCAAATACACTACACAC 526	Db	185 GTAGACTATTTTACATAAATAGTCTACACGCTTGATAGCTTCCAAATACACTACACAC 244		
QY	241 ATTGAACCTTTTTCAGTCAAAAAAGTAGCTGCGGAGTCAAGTAGCGCGCCTTATC 300	QY	241 ATTGAACCTTTTTCAGTCAAAAAAGTAGCTGCGGAGTCAAGTAGCGCGCCTTATC 300		
Db	525 ATTGAACCTTTTTCAGTCAAAAAAGTAGCTGCGGAGTCAAGTAGCGCGCCTTATC 466	Db	245 ATTGAACCTTTTTCAGTCAAAAAAGTAGCTGCGGAGTCAAGTAGCGCGCCTTATC 304		
QY	301 GGGTCGCGCTCTGTCAACGATCAATATTCGACCGGAGTGTGTTCTTATCGT 360	QY	301 GGGTCGCGCTCTGTCAACGATCAATATTCGACCGGAGTGTGTTCTTATCGT 360		
Db	465 GGGTCGCGCTCTGTCAACGATCAATATTCGACCGGAGTGTGTTCTTATCGT 406	Db	305 GGGTCGCGCTCTGTCAACGATCAATATTCGACCGGAGTGTGTTCTTATCGT 364		
QY	361 GACAGAGCGCAGCTTCCTGTGTGTAAACGCGGAGCGCAACTCTTATCGGAACA 420	QY	361 GACAGAGCGCAGCTTCCTGTGTGTAAACGCGGAGCGCAACTCTTATCGGAACA 420		
Db	405 GACAGAGCGCAGCTTCCTGTGTGTAAACGCGGAGCGCAACTCTTATCGGAACA 346	Db	365 GACAGAGCGCAGCTTCCTGTGTGTAAACGCGGAGCGCAACTCTTATCGGAACA 424		
QY	421 GGACGCGCTCCATATCAGCGCGGTATCTCATGCGGTGACCGGAGCGCGCC 480	QY	421 GGACGCGCTCCATATCAGCGCGGTATCTCATGCGGTGACCGGAGCGCGCC 480		
Db	345 GGACGCGCTCCATATCAGCGCGGTATCTCATGCGGTGACCGGAGCGCGCC 286	Db	425 GGACGCGCTCCATATCAGCGCGGTATCTCATGCGGTGACCGGAGCGCGCC 484		
QY	481 GTCCCGCTTATCGCGCTTAAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAG 540	QY	481 GTCCCGCTTATCGCGCTTAAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAG 540		
Db	285 GTCCCGCTTATCGCGCTTAAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAG 226	Db	485 GTCCCGCTTATCGCGCTTAAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAG 544		
QY	541 CATCTGTTACAGCAGCAACAATG 564				
Db	225 CATCTGTTACAGCAGCAACAATG 202				
RESULT 5					
AX766573					
LOCUS	AX766573	2773 bp	DNA	linear	PAT 25-JUN-2003
DEFINITION	Sequence 60 from Patent WO03042244.				
ACCESSION	AX766573				
VERSION	AX766573.1 GI:32260450				
KEYWORDS	synthetic construct				
SOURCE	synthetic construct				
ORGANISM	other sequences; artificial sequences.				
REFERENCE					
1	Klysner,S., Nielsen,F.S., Bratt,T., Voldborg,B. and Mouritsen,S.				
AUTHORS	Novel immunogenic mimetics of multimer proteins				
TITLE	Patent: WO 03042244-A 60 22-MAY-2003;				
JOURNAL	Pharmexa A/S (DK) ; Klysner, Steen (DK) ; Nielsen, Finn Stausholm				

```

Qy      541 CATCTGTT 548
Db      545 CATCTGTT 552

RESULT 6
LOCUS   AY043265/c
DEFINITION Epiphyas postvittana nucleopolyhedrovirus, complete genome.
ACCESSION AY043265
VERSION  AY043265.1 GI:15213125
KEYWORDS
SOURCE  Epiphyas postvittana nucleopolyhedrovirus
ORGANISM Epiphyas postvittana nucleopolyhedrovirus
          Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
          Nucleopolyhedrovirus.
REFERENCE 1 (bases 1 to 118584)
AUTHORS  Hyink,O., Dellow,R.A., Olsen,M., Caradoc-Davies,K.M.B., Drake,K.
          and Ward,V.K.
TITLE    The complete sequence of the Epiphyas postvittana
          nucleopolyhedrovirus genome
JOURNAL  Unpublished
REFERENCE 2 (bases 1 to 118584)
AUTHORS  Hyink,O., Dellow,R.A., Olsen,M., Caradoc-Davies,K.M.B., Drake,K.
          and Ward,V.K.
TITLE    Direct Submission
JOURNAL  Submitted (03-JUN-2001) Microbiology, University of Otago, PO Box
          56, Dunedin, New Zealand
          Location/Qualifiers
FEATURES             source
     1..118584
     /organism="Epiphyas postvittana nucleopolyhedrovirus"
     /mol_type="genomic DNA"
     /db_xref="taxon:70600"
     join(117505..118584,1..609)
     /note="EppomNPV"
     /note="ORF136"
     /codon_start=1
     /product="ORF1629"
     /protein_id="AAK85700.1"
     /db_xref="GI:152131261"
     /translation="MDRQYQSVKSYLIKANNAFANAFATLIAGPEAHNVKQNLIGG
     TVLRNRTAVDLLKLAENIYNDTAYMHVDVTDLNRHYTTLMMRQMLISVTDQHVRLT
     LANIVSRVHELRLYEIVNDVEITTLSDGFVEEYSKYAAROVALSIQMPPEPVTLP
     PPLTPHQVPPPPPLTPLOGLLPVRDVEATPPSTLSKSTLDSFEYFSPNASMIQ
     LPVTPKLIIVKEHLKFKETIISLDPMPATNLLDDQKLPAAPPPTPPPPPLP
     PSNDLPPPPPLLPSPGNVPPPPPIEGMLDMLNIAIAGNNKSEKVTNDARGDMLNL
     IKGVTLKPSKTNKSKDKVDDRADLLNSIKIGVKLPKIKVNTNQLPEAPVTDISVIA
     GTLKKRANISQSSSSSSSSSSQSWDEGNTQTIKANKKHLKAVNLNPFATTHAY
     RTNSELPKLNDVSLDRKPSRVENVNEAKNILDNLKERVKLTNSQLDNAREASLYI
     NDPQFYIQVEDLIFAGRYADAKMHLDAITSGNDELRRLKFKFANDLDVVV"
     complement(606..1343)
     /gene="pol"
     /complement(606..1343)
     /gene="pol"
     /note="ORF1"
     /codon_start=1
     /product="Pol"
     /protein_id="AAK85565.1"
     /db_xref="GI:15213126"
     /translation="MPDYSYRTPTGRTVYDNYKYLKNSGVYKNAKKKHLLEHEDE
     KHLDPDHYMVAEDPFTLPGKNOKLTLPKEIRNVKPTDKLIVNWSGKEFLRETWTFR
     VEDSPFIWNDOEVDVLVINLRPTRNRCYKFLAOKALRWDCDYPVHEVIRIVPEPY
     VGMNNEIRISLAKGGGCPINNIHSETNSFESFVNRVIMENFYIKVIYIGTDSGESE
     EILLESVLVFKVEFADAPLFTGPAY"
     complement(1434..1691)
     /note="ORF2"
     /codon_start=1
     /product="unknown"
     /protein_id="AAK85566.1"
     /db_xref="GI:15213127"
     /translation="MQFELREIYDNYLQRAATAPIMCGTCSETMNVSYQPEHLTTVHKK
     CPLQLSCVRCDDQSSWGQFSAGLMEFHYFCLQNYVQNGV"

CDS
     complement(1713..2090)
     /note="ORF3"
     /codon_start=1
     /product="unknown"
     /protein_id="AAK85567.1"
     /db_xref="GI:15213128"
     /translation="WGSQDSEQKLLKLYCWSQSQTGAGLNDKDLHCLYDLERFVGA
     HLNQADEKAKKKSASAKALRVEIAADRMLAAATPACTDGRWSTLSQAQLDDI
     AREKDIVRIYHLQKQDRLLKQ"
     complement(2104..2715)
     /gene="lef-2"
     complement(2104..2715)
     /gene="lef-2"
     /note="ORF4"
     /codon_start=1
     /product="LEF-2"
     /protein_id="AAK85568.1"
     /db_xref="GI:15213129"
     /translation="MDNVAVNPSANTATIKKNETYLDPNDFMDVLTLSPYTVPRRG
     OLVEISGMRFLPALTMSKPTPEKQFARRSKRNICLKACDGSQTLHKSGLGLKPLC
     IAKIMAEISNASVPRGMYRRKRFNCTYLGNTCTCKKSACLIALHLFYKMDAKCV
     GEVTHLLIKSQDYVKPNSCAKMSKMLCPQANMCKGLNPICNF"
     complement(2717..3070)
     /note="ORF5"
     /codon_start=1
     /product="unknown"
     /protein_id="AAK85569.1"
     /db_xref="GI:15213130"
     /translation="MIYLYYIFFVFPFRESRTSKSPRISHIKLFLFYATLHYNNRR
     INNAPVMVAGHDYDREQLKRDINLSRHSVHELCKRSTTGFCNRLLESSTDTAKPTVV
     IKTTAARQHTICDKL"
     2990..3424
     /note="ORF6"
     /codon_start=1
     /product="unknown"
     /protein_id="AAK85570.1"
     /db_xref="GI:15213131"
     /translation="MADARPLRSTRLTEYKENVVAQIDHLTNLRALINAKTTIANV
     RRGFDNRNAPFAACMDVI VHTYASSDDKITLQPEQLYFRVCRFCNEIADVPDDHDS
     ITRYLCAACGTCLVVDNPLVFTEDDGVQRFMEIQRINAGSEP"
     complement(3426..3935)
     /gene="ptp-1"
     complement(3426..3935)
     /gene="ptp-1"
     /note="ORF7"
     /codon_start=1
     /product="PTP-1"
     /protein_id="AAK85571.1"
     /db_xref="GI:15213132"
     /translation="MPFDRHDTVCNVIETGNTFCFKVPLREMEFYVNDLDRWT
     IKNLIKQCSIGAGIDUTNTPRYDNNMRDEGLLYKRIYVPGVLPDENVVQLFDDV
     VKHPTARCPGMLIGVHCTHGLNRTGYLVCRYMILNISPQDAIARFETARGHKIERQ
     NYIQHLITL"
     4168..5145
     /note="ORF8"
     /codon_start=1
     /product="unknown"
     /protein_id="AAK85572.1"
     /db_xref="GI:15213133"
     /translation="MSLSGKLVVYDYGSMCGVNEQYASVYHLYKIYDEHITNSYVG
     DVPMHREIDATQILSGCALCVNDACHMLDNDTKERLATWYKCGNTKGVCHVYVVL
     KADALVPDKRSGQVDVDTDTLLNTKLPVYLHQYLRVYVYQKLEHVDVEN
     PKSDTGWNRKFCVLTMYRINKSVFASLLTLQKAVVEIQPNAGDKYNCAEVMA
     NVYGRFCGIQKHQFAHKTAMHMFKSMRGQTTMVDERYKSFVIVDFRSHCKETYL
     DLKSHVDSLXYINCTDKQKNAFMDLLCCQSEQLDVCYDIYVNMFLFK"
     complement(5163..5516)
     /note="ORF9; unique to EppomNPV"
     /codon_start=1
     /product="unknown"
     /protein_id="AAK85573.1"
     /db_xref="GI:15213134"
     /translation="MSSNFMVDVNDTHENVETNVKDTLLKPINVWNSDPSSQARCEK
     GAPPVYFAKCVPLESPSRNGVQKIDPTLPALQKITTDITANKTPVLASVKPNNYQDKN

```

Db 114732 CTAGGAATCGTAGACTATTATTTAACTTGAATAGTCTACAGTCTTCTATACGCTCTCTAATACA 114732

Qy 232 CTACCACACATTTGAACCTTTTTCAGTCGCAAAAAGTACGTGT 274

Db 114672 CTACTACACATTTGAATTTTTCAGTCGCAAAAATTTACATAT 114630

RESULT 7

LOCUS AY145471/c 131526 bp DNA circular VRL 07-JUL-2003

DEFINITION Rachiplusia ou multiple nucleopolyhedrovirus, complete genome.

ACCESSION AY145471

VERSION AY145471.1 GI:23476465

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

enhancer

CDS

CDS

CDS

CDS

every Match

st Local Similarity 19.8%; Score 111.8; DB 14; Length 118584;

tches 131; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

112 AACCTTTTTCGACTGCAAAAACACGCTTTTTCAGCGGGGCCCATACATAGTACAACT 171

114792 AACCTTTTTCGACTGCAAAAAGTTTCATTTTCTGAGACCTTATACAGTACAAAA 114733

172 CTACGTTTCGTAGACTATTTTATACAAATAGTCTACACCGTTGTATACGCTCCAAATACA 231

Db 129941 CTACAAATCGTAGACTATTTTATTAGAAATAGTCTACACTGTACAATACGCTCCCAATATA 129882
|||||
Qy 232 CTACACACATGCAACTTTTTCAGTGCAAAAAAGTAGCT 272
|||||
Db 129881 CTACTACACTATCACTTTTTCGATTACAAAAAGTTCAAT 129841
|||||

RESULT 8
NPHE38
LOCUS NPHE38/c
DEFINITION Autographa californica nuclear polyhedrosis virus major early protein (PE-38) gene, complete cds.
ACCESSION M62488.1 GI:332470
VERSION M62488
KEYWORDS major early protein PE-38.
SOURCE Autographa californica nucleopolyhedrovirus
ORGANISM Autographa californica nucleopolyhedrovirus
Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Nucleopolyhedrovirus.
REFERENCE 1 (bases 1 to 1511)
AUTHORS Krappa,R. and Knebel-Morsdorf,D.
TITLE Identification of the very early transcribed baculovirus gene PE-38
JOURNAL J. Virol. 65 (2), 805-812 (1991)
MEDLINE 91101290
PUBMED 1987375
COMMENT Original source text: Autographa californica nuclear polyhedrosis virus DNA, and cDNA to mRNA, isolated from Spodoptera frugiperda.
FEATURES
source
Location/Qualifiers
1..1511
/organism="Autographa californica nucleopolyhedrovirus"
/mol_type="genomic DNA"
/db_xref="taxon:46015"
complement(1..294)
/gene="IE-N"
/genes="IE-N"
complement(1..294)
/genes="IE-N"
352..1453
/genes="PE-38"
352..357
/genes="PE-38"
379..1453
/genes="PE-38"
/product="immediate-early protein PE-38"
428..1393
/gene="PE-38"
/codon_start=1
/product="immediate-early protein PE-38"
/protein_id="AAA46733.1"
/db_xref="GI:332471"
/translations="MPRDTNNRHSPTVERPTLEDLRQLQDNLDSINRRDRMORQOE
ENLRVQRQRQQLRSIQMEQRMALNPEVINFPECSVCLETYQSQNDTCP
FLIPTCDGFCFKVINIQSNANNIPHSVCCPLCNQVMKRSUKPNAAVTVCKEYK
KTOERVPPVQYKNIKIVQERSVISEDNDNDNMENQAKIALEAELEBKNS
DQVSNRLIENLNEQIQELQVRLTVPQGITVQIQIGRDSAPAEINERFR
SLVYSITSELFIENRVHSIQYVYAGTSASSCDVAVTVNFGFN"
1426..1431
/genes="PE-38"
1433..1438
/genes="PE-38"

polyA_signal
polyA_signal

ORIGIN
Query Match 18.0%; Score 101.4; DB 14; Length 1511;
Best Local Similarity 75.4%; Pred. No. 4.4e-21;
Matches 126; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 112 AACCTTTTTCAGTCAAAAAACACGCTTTTTCAGCGGGCCATACATAGTACAAACT 171
|||||
Db 286 AATTTTTCATGCAAAAAGTTTACCTTTTGCCTGACACTCCATATACATGACATCT 227
|||||

Qy 172 CTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACA 231
|||||
Db 226 CTACAAATCGTAGACTATTTTATTAGAAATAGTCTACACTGTACGATCGTCCCAATATA 167
|||||

Qy 232 CTACACACATGCAACTTTTTCAGTCAAAAAAGTAGTGTGCGC 278
|||||
Db 166 CTACTACACTATCACTTTTTCGATTACAAAAAGTTTCATTTTTCG 120
|||||

RESULT 9
A48542/c
LOCUS A48542
DEFINITION Sequence 1 from Patent WO9601320.
ACCESSION A48542
VERSION A48542.1 GI:2302312
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 133894)
AUTHORS Bishop,D., Possee,R. and Ayres,M.
TITLE AUTOGRAHA CALIFORNICA COMPLETE GENOME SEQUENCE
JOURNAL Patent: WO 9601320-A 1 18-JAN-1996;
NATURAL ENVIRONMENT RES (GB)
COMMENT Other publication AU 2897295 960125.
FEATURES
source
1..133894
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 18.0%; Score 101.4; DB 6; Length 133894;
Best Local Similarity 75.4%; Pred. No. 1.2e-20;
Matches 126; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 112 AACCTTTTTCAGTCAAAAAACACGCTTTTTCAGCGGGCCATACATAGTACAAACT 171
|||||
Db 132384 AATTTTTCGCAATGCAAAAAGTTTCACCTTTTGCCTGACACTCCATATACATGACATCT 132325
|||||

Qy 172 CTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACA 231
|||||
Db 132324 CTACAAATCGTAGACTATTTTATTAGATAGTCTACACTGTACGATCGTCCCAATATA 132265
|||||

Qy 232 CTACACACATGCAACTTTTTCAGTCAAAAAAGTAGTGTGCGC 278
|||||
Db 132264 CTACTACACTATCACTTTTTCGATTACAAAAAGTTTCATTTTTCG 132218
|||||

RESULT 10
L22858/c
LOCUS L22858
DEFINITION Autographa californica nucleopolyhedrovirus clone C6, complete genome.
ACCESSION L22858
VERSION L22858.1 GI:510708
KEYWORDS
SOURCE Autographa californica nucleopolyhedrovirus
ORGANISM Autographa californica nucleopolyhedrovirus
Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Nucleopolyhedrovirus.
REFERENCE 1 (bases 1 to 133894)
AUTHORS Ayres,M.D., Howard,S.C., Kuzio,J., Lopez-Ferber,M. and Possee,R.D.
TITLE The complete DNA sequence of Autographa californica nuclear polyhedrosis virus
JOURNAL Virology 202 (2), 586-605 (1994)
MEDLINE 94303173
PUBMED 8030224
REFERENCE 2 (bases 1 to 133894)
AUTHORS Kuzio,J.
TITLE Direct Submission
JOURNAL Submitted (08-MAR-1999) NCBI, Bethesda, MD 20894, USA
FEATURES
source
Location/Qualifiers
1..133894
/organism="Autographa californica nucleopolyhedrovirus"
/mol_type="genomic DNA"

```

/db_xref="taxon:46015"
/clone="C6"
join(133883..133894,1..445)
/standard_name="hri"
/note="5 copies of 30 bp imperfect palindromic sequence;
the EcoRI site in the first palindrome is at residue 1 of
the linearized genome"
/function="enhancer; replication origin"
/rpt_type=dispersed
503..1009
/gene="Ac-ptp"
/note="synonym: AcOrf-1"
503..1009
/gene="Ac-ptp"
/note="PTP; 19288 Da primary translation product"
/codon_start=1
/product="protein tyrosine phosphatase"
/protein_id="AAA66631.1"
/db_xref="GI:559070"
/translation="MFPAWHNYLQCGOVIKDSNLICPKTLPRLPELAYVTSEEDVMT
AEQIVQNPSGAIDLTNTSKYIDGVHFLAGLLYKKIQVGGTLPESIVQEFIDT
VKEFTKCPGMLGVHCTHGINRTGYMCRYVLMHTLGIAPQEAIDRFKARGHKIERQ
NYVQDLTI"
complement(1041..2027)
/gene="Ac-bro"
/note="synonym: AcOrf-2"
complement(1041..2027)
/gene="Ac-bro"
/note="BRO; 37769 Da primary translation product"
/codon_start=1
/product="baculovirus repeated ORF"
/protein_id="AAA66632.1"
/db_xref="GI:559071"
/translation="MARVIGEFKGGDTFNLRLVLRDQOVRFVAKDVANSKYTVC
DKLRVHVDNKYLFQETIQNGGTSNVVKGDPYLPQHTVLITKSGVIQIMKS
KLPALEQLLEBEVQVLCQKYDPAIKQREESKQLVTKLIATPTEHTNALQAV
VAQTEELVQKEFTIRVAIKQIEAKDLQVTRVMTDLNRMVTGFOETWQKKDEIM
OKKQAOVTLVAKVVDLSRAVOYPADKKEHPLCVTDGDTFTTALIGOKTYVENQKH
KRNINVAIVENTRNPTVDWNNATDLQAKRSIVLVRKRNKLIG"
/gene="Ac-ctx"
/note="synonym: AcOrf-3"
complement(2084..2245)
/gene="Ac-ctx"
/note="CTX; 5590 Da primary translation product"
/codon_start=1
/product="conotoxin-like peptide"
/protein_id="AAA66633.1"
/db_xref="GI:559072"
/translation="WQIKTVLLAFAMFAALNAQHVAACAEFGAVCVHNDCCSGACS
PIFNVCPLPQ"
complement(2295..2750)
/gene="AcOrf-4"
complement(2295..2750)
/gene="AcOrf-4"
/note="17577 Da primary translation product"
/codon_start=1
/product="AcOrf-4 peptide"
/protein_id="AAA66634.1"
/db_xref="GI:559073"
/translation="WKLTVMASLLKYALRLTRYKENIIPHOHLVRLRLDLMGKIK
SDHVRNTRNRDLISACMOINVRYTPNATIDMRKOPNCIYFRIQYCHLEADVPS
PDDHSVRYLVCAGTPIVDHPLDVFCHTEGYNELLEVRQVAGGEL"
2779..3108
/gene="AcOrf-5"
2779..3108
/gene="AcOrf-5"
/note="12435 Da primary translation product"
/codon_start=1
/product="AcOrf-5 peptide"
/protein_id="AAA66635.1"
/db_xref="GI:559074"
/translation="MYRTSRINNAFVAVSQHDYDRDQIKRELSLRNVHDLCTRSGT

```

```

SFDCNKELRSDMTVPVTTITPKEADYKITEYVDVKTIKPSNRPLVESGPLVREAA
KYGSCIV"
/3089..3721
/note="synonym: AcOrf-6"
3089..3721
/gene="Ac-lef2"
/note="LEF2; 23926 Da primary translation product"
/codon_start=1
/product="late expression factor 2"
/protein_id="AAA66636.1"
/db_xref="GI:559075"
/translation="MANASYNVWSPILRASCLDKKATVLDIPDDFIDKLTLTPYTVFY
NGVLVKISGLRLYMLLTAPTINEIKSNFKSKSKNI CMKECVGKXNVDMLNKX
INMPCKIKIILNDLKENNVPRGMYRRKFIILNCYIANVVSACAKENCLIKALTHIFYN
HDSKCVGEVHLLIKSQDVYKPCQKMTVDKLCPPFAGNCKGLNPICNY"
complement(3759..4364)
/gene="Ac-ORF603"
/note="synonym: AcOrf-7"
complement(3759..4364)
/gene="Ac-ORF603"
/note="ORF603 peptide; 23612 Da primary translation
product"
/codon_start=1
/protein_id="AAA66637.1"
/db_xref="GI:559076"
/translation="MAVIFNNKQLADNSIEKGGLFLFNGSYNILESYNPVLKNG
VIELEAAAYAGNLIYKTDPPKFDYINLIKATHSELPENSTVNVYRKTRMSGTHI
PKKDIYIDNKKFTLYDYIDNNYVNFYEKEKEKEEEDDKASSLCENKII
LSQINCSEFENDFKYLSYDNYAFSIIDNTNVLVAFGLYR"
4520..5257
/gene="Ac-PH"
/note="synonym: AcOrf-8"
4520..5257
/gene="Ac-PH"
/note="polyhedrin; 28642 Da primary translation product"
/codon_start=1
/product="major occlusion body protein"
/protein_id="AAA66638.1"
/db_xref="GI:559077"
/translation="MPDYSYRPTIGRTYVDNKKYKNGAVIKNAKKKGFPAEHEIEE
ATLDPLNDLYVAEDPFLGPGKQKLIKFKEIRNVKPDTKLVGKGGKEFVETWTRF
MDSFPITVNDQEVMDVFLVNMVRPNRCVYKFLAQHALRCDPDYVHDVIRIVEPSW
VGSNNVYRISLAKKGGCCPIMNLHSEYTSFEQFIDRVIMENFYKPIVYIGTDSAEIE
EILLESVLVKVEEPADAPLFTGPAY"
complement(5287..6918)
/gene="Ac-ORF1629"
/note="synonym: AcOrf-9; Ac-61k; Ac-vp78"
complement(5287..6918)
/gene="Ac-ORF1629"
/note="61k; vp78; required for virus replication; 60713 Da
primary translation product"
/codon_start=1
/product="viral capsid associated protein"
/protein_id="AAA66639.1"
/db_xref="GI:559078"
/translation="MTNRYVESVQSYLFNNRNKKIDAHQPFERVDTAEAQILKKNIYD
NTVVLNRDVLILKLANDFNKAYMVDSEVSRHYNVAVKMKRLVIGVDFSLRQ
SLYNTIAYIERLNLIGTNDSEITMLIADFYDLYSNYNIELPPQALPRSRPSVQP
AAPAPVTIRVREQTKPQIIIPAPPPPPSPVNIIPAPPPPPSMSELPAPMPPTPEP
QPAAPLDRQOLLEAKRNRKLRVPKPTAPETSTIVEVTVLPKETPEPKPPSA
SPPPPPPPPPAPPAIPPMVDLSSAPPPLVDLPSEMLPPPPAPPSLSNLSKSGT
VRLPKAQKRCQSEIIPKSSNTNLIADVLADTINRRRVAMAKSSSEATSNDBGDDON
RPNKANTPVKIYQALFNVTSSQLYTNDSDEKNTKHNILNDVEPLLQNTQTNIDK
ARLLQDLASFSVALSENPSPGPAIGSEKQPLFETNRNLFYKSIEDLIFKPKYKOAENH
LIFALTYHPKDYFNEKLLKYVQLSVNQQRTESSA"
6917..7735
/gene="Ac-pk-1"
/note="synonym: AcOrf-10"
6917..7735
/gene="Ac-pk-1"
/note="PK1; 31978 Da primary translation product"
/codon_start=1

```


/db_xref="GI:30269988"
 /translation="MHSVRNLFDSNMSSLLVAVYGAFLNLLHEKGYHLYRIV
 HEHLTETVSNASCVRRDIATAPCFENGFCNLFARQLDVTDAARLAANRQDKTG
 LCAVOLALAEIDKYAPLEKESVIGNIFALDTIADIPSNALDDPQTIIYSGPKDFVD
 MNLAHVADFPDPKIXAEGWYKFCVLTYMERLTVNAVPTLMTLRLQDAVIRKQV
 POKNGCANPANYVFCGIGRKHFSQKAAASYILFYQYRNLLTPKDERHPSFGVI
 KDFGRCKETTYDLRAEADLLYINATTDKKNALFDLLCCVNAADLDVDCYDIYDNF
 YFICK"
 complement (5854. .6387)

CDS

Query Match 16.3%; Score 92.2; DB 14; Length 129609;
 Best Local Similarity 78.3%; Pred. No. 1.1e-17;
 Matches 123; Conservative 0; Mismatches 33; Indels 1; Gaps 1;
 QY 112 AACCTTTTGGCACTCAAAAACACGCTTTTGCACGGGCGCCCATACATAGTACAAACT 171
 DB 124891 AACCTTTTGGAGTGCAGAAATA-ATCATTAATTTCAAGAGCCCATATACAGAACAAAGT 124833
 QY 172 CTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACA 231
 DB 124832 CTTCGTTTCGTAGACTATTTTACTTAATAGTCTACACTGTTTAATATGCTCTCTATACA 124773
 QY 232 CTACACACATGAACCTTTTTCGAGTGCAGAAAAGT 268
 DB 124772 CTACACGTAATGTACATTTTTCGACTGCAAAAAGT 124736

RESULT 15

NPBBMIEN 2178 bp DNA linear VRL 03-FEB-1999
 LOCUS Bombyx mori nuclear polyhedrosis virus genes for BmIE-N and BmPE36,
 DEFINITION partial cds and complete cds.
 ACCESSION D14467
 VERSION D14467.1 GI:222184
 KEYWORDS BmIE-N; BmPE36; immediate-early gene.
 SOURCE Bombyx mori nucleopolyhedrovirus
 ORGANISM Bombyx mori nucleopolyhedrovirus
 Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 Nucleopolyhedrovirus.
 1 (bases 1 to 2178)
 Yamanoto,G. and Kobayashi,M.
 Nucleotide sequence and transcriptional analysis of an
 immediate-early gene, PE36, from Bombyx mori nuclear polyhedrosis
 virus
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2178)
 Yamanoto,G.
 AUTHORS Direct Submission
 TITLE Submitted (18-FEB-1993) Gentaro Yamamoto, Nagoya University Faculty
 JOURNAL of Agriculture, Dept. of Agricultural Chemistry; Nagoya, Aichi
 464-01, Japan (Tel:052-781-5111 (ex.6238), Fax:052-781-4447)
 COMMENT Submitted (18-FEB-1993) to DDBJ by:
 Gentaro Yamamoto
 Faculty of Agriculture
 Nagoya University
 Nagoya, Aichi 464-01
 Japan
 Phone: 052-781-5111
 Fax: 052-781-4447.
 Location/Qualifiers
 1..2178
 /organism="Bombyx mori nucleopolyhedrovirus"
 /mol_type="genomic DNA"
 /isolate="BmNPV-NS"
 /db_xref="taxon:271108"
 complement (<1. .414)
 /codon_start=1
 /product="BmIE-N"
 /protein_id="BAA03363.1"
 /db_xref="GI:222185"
 /translation="MSRQINAVTSSSSSRHRLSLRRINFTTSPGALPYSISR
 SQPVSSRRRRQSRQVRSDNVQIIGNAPLTTHTHSQGVTHVHGQVNIISND
 DPLLSQEDDTIESVDRLSLSTRIRRAITRAALQ"

FEATURES
source

CDS

TATA_signal
 TATA_signal
 CDS

complement (487. .493)
 782. .788
 847. .1776
 /codon_start=1
 /product="BmPE36"
 /protein_id="BAA03364.1"
 /db_xref="GI:222186"
 /translation="MPRETNRRRRSTPYERPTLEDLHROLEDALERRRYOTRRRQ
 NLRITIOIKQORMAELKKRPVINFKECSVCPEYQQSNDTCPELPTTCDHGFCF
 KXVIDLQSNAMNIPHSIVCCPLCNQVQKMRSLAPNAVVTCKFKYKQERVPAVQOYK
 NIIVLQERSVISVENNDKCDINENQAKIVALEELKTEKNHSDQVTSNRQLTEE
 NTRLNEQVQLQQRVTLRPFQGITVNPQIGRDDSPAPALNFRSLAYSITSLFIE
 NRHSIQNYVYAGTSGAASSCDVNVTVNFGFEN"
 1827. .1832

polya_signal
 ORIGIN

Query Match 16.3%; Score 92; DB 14; Length 2178;
 Best Local Similarity 72.6%; Pred. No. 5.3e-18;
 Matches 119; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
 QY 115 TTTTTCGCACTGCAAAAACACACGCTTTTGCACGGGCGCCCATACATAGTACAACTCTA 174
 DB 712 TTTTTCGCAATACAAAAGAGTTTCGCTATGTTTGACATTTAATATACAGTACGAACTCTA 653
 QY 175 CGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATAGCTCCAAATACACTA 234
 DB 652 CAAATCGTAGACTATTTTATTAGATAAGTCTACACTGTACTATACGCTCTCAATATACTA 593
 QY 235 CCACACATTGAACCTTTTTCGAGTGCAGAAAAGTACGTGTCGGC 278
 DB 592 CTACACTATCAACTTTTTCGATTTACAAAAAGTTTCATTTTTCG 549

Search completed: May 10, 2005, 05:52:38
 Job time : 3422.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 01:20:43 ; Search time 616.764 Seconds
(without alignments)
5413.313 Million cell updates/sec

Title: US-09-896-888A-1
Perfect score: 564
Sequence: 1 catgatgataaaacatgtat.....tggtacagcgacacacatg 564

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	564	100.0	564	2	AAV62487
2	548	97.2	560	12	ADQ48575
3	548	97.2	2773	9	AAL61306
4	546.4	96.9	5038	12	ADQ48539
5	101.4	18.0	279	2	AAT13730
6	101.4	18.0	13894	2	AAT13635
7	93.4	16.6	28413	10	ADC51646
8	62	11.0	141	12	ADQ48576
9	39.6	7.0	6289	8	ABZ10059
10	39.6	7.0	9289	4	AAS46501
11	39.6	7.0	9289	10	AD884121
12	38	6.7	6289	8	ABZ10205
13	38	6.7	9289	10	AD884197
14	36.2	6.4	2000	8	ADA71938
15	35.4	6.3	986	6	ABQ68715
16	35.4	6.3	1549	6	ABQ70339
17	34.4	6.1	855	2	AAV34232
18	34.4	6.1	855	8	ACD08103
19	34	6.0	18977	4	ABL19380
20	33.8	6.0	1991	6	AAS18541

C 21	33.8	6.0	2117	11	ACN44823	ACN44823 Human mRN
C 22	33.8	6.0	3056	8	ABZ42667	ABZ42667 Human gal
C 23	33.8	6.0	3083	2	AAQ97304	AAQ97304 Galanin r
C 24	33.8	6.0	3203	13	ACN43223	ACN43223 Human dia
C 25	33.8	6.0	39344	11	ACN44822	ACN44822 Human gen
C 26	33.2	5.9	43053	8	ABZ73951	ABZ73951 Secreted
C 27	33.2	5.9	43053	10	ABZ67535	ABZ67535 Human sec
C 28	33.2	5.9	43056	8	ABZ73950	ABZ73950 Secreted
C 29	33.2	5.9	43056	10	ABZ67534	ABZ67534 Human sec
C 30	33	5.9	14429	6	ABL34242	ABL34242 Human imm
C 31	33	5.9	14429	6	ABQ67097	ABQ67097 Human ang
C 32	32.8	5.8	426	3	AH30400	AH30400 Human col
C 33	32.8	5.8	454	6	ABL93815	ABL93815 Arabidops
C 34	32.8	5.8	1110	3	AAC37066	AAC37066 Arabidops
C 35	32.8	5.8	1251	3	AAC36504	AAC36504 Arabidops
C 36	32.8	5.8	2000	6	ABZ17343	ABZ17343 Arabidops
C 37	32.8	5.8	2000	8	ADA68882	ADA68882 Arabidops
C 38	32.8	5.8	2780	8	ADA70711	ADA70711 Rice gene
C 39	32.6	5.8	495	5	ADL37609	ADL37609 Human ova
C 40	32.6	5.8	60604	12	ADO48003	ADO48003 Human HIP
C 41	32.6	5.7	1268	4	ABL18801	ABL18801 Drosophil
C 42	32.4	5.7	1342	4	ABL27503	ABL27503 Drosophil
C 43	32.4	5.7	3448	4	ABL27502	ABL27502 Drosophil
C 44	32.4	5.7	3600	4	ABL18800	ABL18800 Drosophil

ALIGNMENTS

RESULT 1
AAV62487
ID AAV62487 standard; DNA; 564 BP.

XX AAV62487;

AC AAV62487;

DT 17-OCT-2003 (revised)

DT 19-JAN-1999 (first entry)

XX O. pseudotsugata multicapsid nucleopolyhedrosis virus ie2 promoter.

XX Orgyia pseudotsugata; multicapsid; nucleopolyhedrosis virus; OpMNPV;

XX Op ie2; promoter; shuttle vector; transformation; melanotransferrin;

XX immediate early baculovirus promoter; prokaryotic; transposon;

XX bleomycin/phleomycin-type antibiotic; insect cell; transposon;

XX ion transport peptide hormone; ss.

XX Orgyia pseudotsugata; polyhedrosis virus.

XX WO9844141-A2.

XX 08-OCT-1998.

XX 26-MAR-1998; 98WO-CA000282.

XX 27-MAR-1997; 97US-0049946P.

XX 28-JAN-1998; 98CA-02221819.

XX (UYER-) UNIV BRITISH COLUMBIA.

XX Grigliatti TA, Theilmann DA, Pfeifer TA, Hegedus DD;

XX WPI; 1998-557129/47.

XX Expression vectors for transforming insect cells from disparate lines -
PT useful to express heterologous DNA, e.g. to allow study of gene
PT expression and produce commercially important proteins.
PS Claim 10; Page 82; 121pp; English.
XX This represents a Orgyia pseudotsugata multicapsid nucleopolyhedrosis
CC virus (OpMNPV) immediate early 2 (Op ie2) promoter sequence. The
CC invention provides a new shuttle vector for transforming insect cells

CC that comprises: (i) prokaryotic origin of replication; (ii) insect
CC promoter having homology to, and capable of functioning as, an immediate
CC early baculovirus promoter; (iii) prokaryotic promoter sequence; and (iv)
CC selectable marker capable of conferring resistance to a bleomycin/
CC phleomycin-type antibiotic under transcriptional control of (ii) and
CC (iii), in insect and prokaryotic cells respectively. The vectors can be
CC used to stably transform (especially insect) cells with heterologous DNA,
CC useful to allow study of gene expression and direct expression of
CC heterologous gene products e.g. commercially important proteins. They are
CC especially useful to allow expression of the heterologous
CC melanotransferrins, ion transport peptide hormones or biologically active
CC derivatives in insect cells. They enable transformation of insect cell
CC lines from disparate species, allowing screening of lines for optimum
CC post-translational modification of particular proteins. Shuttle vectors
CC further comprising DNA transposable elements defining a transposon can be
CC used to optimise heterologous protein expression and facilitate selection
CC of desired transformants. (updated on 17-OCT-2003 to standardise OS
CC field)
XX
SQ Sequence 564 BP; 147 A; 160 C; 116 G; 141 T; 0 U; 0 Other;

Query Match 100.0%; Score 564; DB 2; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.1e-172;
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGATGATAAACAATGATGCTGCTAATGTTGCTTCAACAACAACTTCTGTTGAACCTGTG 60
DB 1 CATGATGATAAACAATGATGCTGCTAATGTTGCTTCAACAACAACTTCTGTTGAACCTGTG 60
QY 61 TTTTCATGTTGGCCAAAGCAAGCACTTTATATCTCGGTGGCTCCCAACCACTTTTTT 120
DB 61 TTTTCATGTTGGCCAAAGCAAGCACTTTATATCTCGGTGGCTCCCAACCACTTTTTT 120
QY 121 GCATGCAAAAACACAGCTTTTTCACGGGGCCCATACATAGTACAACTCTAGTTTC 180
DB 121 GCATGCAAAAACACAGCTTTTTCACGGGGCCCATACATAGTACAACTCTAGTTTC 180
QY 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCACAC 240
DB 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCACAC 240
QY 241 ATTGAACCTTTTTCAGTGCAAAAAGTACGTGCGGAGTCAGTACGTAGCGGGCTTATC 300
DB 241 ATTGAACCTTTTTCAGTGCAAAAAGTACGTGCGGAGTCAGTACGTAGCGGGCTTATC 300
QY 301 GGGTCGCTCTGTACCTAGCAATCATATTCGACCGGACGAGTCTTCTTATCGT 360
DB 301 GGGTCGCTCTGTACCTAGCAATCATATTCGACCGGACGAGTCTTCTTATCGT 360
QY 361 GACAGGCGCCGAGCTTCTGTGTTGCTAACCGGACCGGACCACTCTTATCGGAACA 420
DB 361 GACAGGCGCCGAGCTTCTGTGTTGCTAACCGGACCGGACCACTCTTATCGGAACA 420
QY 421 GGACGCGCTCATATACAGCGCGGTTTATCTCATGCGCGTGACCGGACAGGCGGCC 480
DB 421 GGACGCGCTCATATACAGCGCGGTTTATCTCATGCGCGTGACCGGACAGGCGGCC 480
QY 481 GTCCGCTTATCGCGCTTAAATACAGCGCGGACGATCTGTTAAACACAGTTGACAG 540
DB 481 GTCCGCTTATCGCGCTTAAATACAGCGCGGACGATCTGTTAAACACAGTTGACAG 540
QY 541 CATCTGTTTACAGCGCAACAATG 564
DB 541 CATCTGTTTACAGCGCAACAATG 564

RESULT 2
ADQ48575
ID ADQ48575 standard; DNA; 560 BP.
XX
AC ADQ48575;
XX
DT 09-SEP-2004 (first entry)

XX OpiE2 promoter DNA sequence.
DE
XX
KW viral vector; recombination site; recombinant virus;
KW replication-defective particle generation; gene expression inhibition;
KW gene therapy vector; ds; OpiE2 promoter.
OS
XX Unidentified.
XX WO2004009768-A2.
XX
XX 29-JAN-2004.
XX
XX 18-JUL-2003; 2003WO-US022437.
XX
XX 18-JUL-2002; 2002US-0396335P.
XX 26-JUL-2002; 2002US-0398617P.
XX 19-NOV-2002; 2002US-0427231P.
XX 24-MAR-2003; 2003US-0456496P.
XX 03-JUN-2003; 2003US-0474940P.
XX
XX (INVI-) INVITROGEN CORP.
XX (BENN/) BENNETT R P.
XX (WELC/) WELCH P J.
XX (HARW/) HARWOOD S.
XX (MADD/) MADDEN K.
XX (FRIM/) FRIMPONG K.
XX (FRAN/) FRANK K E.
XX
XX Bennett RP, Welch PJ, Harwood S, Madden K, Frimpong K, Franke KE;
XX WPI; 2004-132944/13.
XX
XX Novel nucleic acid molecule comprising all or a portion of one or more
XX viral genome and further comprising two or more recombination sites that
XX do not substantially recombine with each other, useful as gene therapy.
XX
XX Disclosure; Fig 16; 555pp; English.
XX
XX The invention comprises a nucleic acid molecule consisting of all or a
XX portion of one or more viral/baculoviral genomes and further containing
XX at least two recombination sites that do not substantially recombine with
XX each other. One or more of the recombination sites is capable of
XX undergoing recombination with a compatible recombination site in the
XX presence of one or more proteins active in lambda recombination. The
XX nucleic acid of the invention replicates in prokaryotic and eukaryotic
XX cells. The nucleic acid of the invention is useful for constructing a
XX recombinant virus, generating replication-defective particles,
XX preventing/inhibiting the expression of one or more genes in an organism,
XX and are useful as gene therapy vectors. The nucleic acid of the invention
XX is also useful for producing and expressing fusion polypeptides. The
XX present DNA sequence represents the OpiE2 promoter that was used in the
XX exemplification of the invention.

Sequence 560 BP; 144 A; 156 C; 116 G; 144 T; 0 U; 0 Other;
Query Match 97.2%; Score 548; DB 12; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.8e-167;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGATGATAAACAATGATGCTGCTAATGTTGCTTCAACAACAACTTCTGTTGAACCTGTG 60
DB 5 CATGATGATAAACAATGATGCTGCTAATGTTGCTTCAACAACAACTTCTGTTGAACCTGTG 64
QY 61 TTTTCATGTTGGCCAAAGCAAGCACTTTTATATCTCGGTGGCTCCCAACCACTTTTTT 120
DB 65 TTTTCATGTTGGCCAAAGCAAGCACTTTTATATCTCGGTGGCTCCCAACCACTTTTTT 124
QY 121 GCATGCAAAAACACAGCTTTTTCACGGGGCCCATACATAGTACAACTCTAGTTTC 180
DB 125 GCATGCAAAAACACAGCTTTTTCACGGGGCCCATACATAGTACAACTCTAGTTTC 184
QY 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCACAC 240

Db 185 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCAAAATACACTACCACAC 244
Qy 241 ATTGAACCTTTTTCAGTGCACAAAAAGTACGTCGCGCAGTCACGTAGGCGCGCTTATC 300
Db 245 ATTGAACCTTTTTCAGTGCACAAAAAGTACGTCGCGCAGTCACGTAGGCGCGCTTATC 304
Qy 301 GGTCGCGTCTCTCAGTACAGTAAGATCACAATATCGGACCGGACGAGTGTCTTATCGT 360
Db 305 GGTCGCGTCTCTCAGTACAGTAAGATCACAATATCGGACCGGACGAGTGTCTTATCGT 364
Qy 361 GACAGACCGCAGCTTCTGTTGCTTAACCGCAGCGGACGCACTCTTATCGAACA 420
Db 365 GACAGACCGCAGCTTCTGTTGCTTAACCGCAGCGGACGCACTCTTATCGAACA 424
Qy 421 GGACGGCGCTTCCATATCAGCGCGCGCTTATCTCATGCGCGTGACCGGACACGAGGCGCC 480
Db 425 GGACGGCGCTTCCATATCAGCGCGCGCTTATCTCATGCGCGTGACCGGACACGAGGCGCC 484
Qy 481 GTCCCGTTATCGCGCTATAAATACAGCCCGCAACGATCTGTAAACACAGTTGAACAG 540
Db 485 GTCCCGTTATCGCGCTATAAATACAGCCCGCAACGATCTGTGTAAACACAGTTGAACAG 544
Qy 541 CATCTGTT 548
Db 545 CATCTGTT 552

RESULT 3
AAL61306

ID AAL61306 standard; DNA; 2773 BP.

XX AC

XX AC

XX AC

XX 22-SEP-2003 (first entry)

XX p2Op2F expression vector for insect cells.

XX Multimeric protein; interleukin 5; IL5; TNFalpha; inflammatory disease;
XX tumour necrosis factor alpha; gene therapy; arthritis; gene; ds.

XX Unidentified.

XX WO2003042244-A2.

XX 22-MAY-2003.

XX 15-NOV-2002; 2002WO-DK000764.

XX 16-NOV-2001; 2001DK-00001702.

XX 16-NOV-2001; 2001US-0331575P.

XX (PHAR-) PHARMEXA AS.

XX (KLYS/) KLYSNER S.

XX (NIEL/) NIELSEN F S.

XX (BRAT/) BRATT T.

XX (VOLD/) VOLDORF B.

XX (MOUR/) MOURITSEN S.

XX Klynsner S, Nielsen FS, Bratt T, Voldborg B, Mouritsen S;
XX WPI; 2003-449558/42.XX New immunogenic analogue of a polymeric protein, useful for preparing a
XX composition for treating inflammatory diseases e.g. arthritis.

XX Disclosure; Page 195-196; 196pp; English.

XX The invention relates to immunogenic analogues of multimeric proteins
XX such as immunogenic variants of interleukin 5 (IL5) and tumour necrosis
XX factor alpha (TNF, TNFalpha) and methods for production of immunogenic
XX analogues. The immunogenic analogue is useful for preparing a composition
XX for treating inflammatory diseases, e.g., arthritis. It is also used in

CC gene therapy. The present sequence is p2Op2F expression vector for
CC insect cells. This sequence is used to illustrate the method of the
CC invention

XX Sequence 2773 BP; 669 A; 750 C; 673 G; 681 T; 0 U; 0 Other;
SQ

Query Match 97.2%; Score 548; DB 9; Length 2773;

Best Local Similarity 100.0%; Pred. No. 3.8e-167;

Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGATGATAAACAATGATGTTGCTTAATGTTCTTCAACAACAATTCGTCTCAACTGTG 60

Db 5 CATGATGATAAACAATGATGTTGCTTAATGTTCTTCAACAACAATTCGTCTCAACTGTG 64

Qy 61 TTTTCATGTTGCAACAAGACACCTTTTATATCGGTGGCTCCCAACCACTCTTTT 120

Db 65 TTTTCATGTTGCAACAAGACACCTTTTATATCGGTGGCTCCCAACCACTCTTTT 124

Qy 121 GCACGCAAAAAACAAGCTTTTTCAGCGGGGCCATATAGTACAAACTCTACGTTTC 180

Db 125 GCACGCAAAAAACAAGCTTTTTCAGCGGGGCCATATAGTACAAACTCTACGTTTC 184

Qy 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCAC 240

Db 185 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCAC 244

Qy 241 ATTGAACCTTTTTCAGTGCACAAAAAGTACGTCGCGCAGTCACGTAGGCGCGCTTATC 300

Db 245 ATTGAACCTTTTTCAGTGCACAAAAAGTACGTCGCGCAGTCACGTAGGCGCGCTTATC 304

Qy 301 GGTCGCGTCTCTCAGTACAGTAAGATCACAATATCGGACCGGACGAGTGTCTTATCGT 360

Db 305 GGTCGCGTCTCTCAGTACAGTAAGATCACAATATCGGACCGGACGAGTGTCTTATCGT 364

Qy 361 GACAGACCGCAGCTTCTGTTGCTTAACCGCAGCGGACGCACTCTTATCGAACA 420

Db 365 GACAGACCGCAGCTTCTGTTGCTTAACCGCAGCGGACGCACTCTTATCGAACA 424

Qy 421 GGACGGCGCTTCCATATCAGCGCGCGCTTATCTCATGCGCGTGACCGGACACGAGGCGCC 480

Db 425 GGACGGCGCTTCCATATCAGCGCGCGCTTATCTCATGCGCGTGACCGGACACGAGGCGCC 484

Qy 481 GTCCCGTTATCGCGCTATAAATACAGCCCGCAACGATCTGTAAACACAGTTGAACAG 540

Db 485 GTCCCGTTATCGCGCTATAAATACAGCCCGCAACGATCTGTGTAAACACAGTTGAACAG 544

Qy 541 CATCTGTT 548

Db 545 CATCTGTT 552

RESULT 4

ADQ48539 standard; DNA; 5038 BP.

XX AC

XX AC

XX ADQ48539;

XX 09-SEP-2004 (first entry)

XX Viral vector-related plasmid - pIB/V5-His-DEST.

XX viral vector; recombination site; recombinant virus;
XX replication-defective particle generation; gene expression inhibition;

XX gene therapy vector; ds; plasmid.

XX Unidentified.

XX WO2004009768-A2.

XX 29-JAN-2004.

XX 18-JUL-2003; 2003WO-US022437.

XX

PR 18-JUL-2002; 2002US-0396335P.
 PR 26-JUL-2002; 2002US-0398617P.
 PR 19-NOV-2002; 2002US-0427231P.
 PR 24-MAR-2003; 2003US-0456496P.
 PR 03-JUN-2003; 2003US-0474940P.
 XX (INVI-) INVITROGEN CORP.
 PA (BENN/) BENNETT R P.
 PA (WELC/) WELCH P J.
 PA (HARW/) HARWOOD S.
 PA (MADD/) MADDEN K.
 PA (FRIM/) FRIMPONG K.
 PA (FRAN/) FRANK K E.
 XX
 PI Bennett RP, Welch PJ, Harwood S, Madden K, Frimpong K, Franke KE;
 XX WPI; 2004-132944/13.
 DR
 XX
 XX Novel nucleic acid molecule comprising all or a portion of one or more
 PT viral genome and further comprising two or more recombination sites that
 PT do not substantially recombine with each other, useful as gene therapy.
 XX
 XX Example 18; Page 395-403; 555pp; English.
 XX
 CC The invention comprises a nucleic acid molecule consisting of all or a
 CC portion of one or more viral/baculoviral genomes and further containing
 CC at least two recombination sites that do not substantially recombine with
 CC each other. One or more of the recombination sites is capable of
 CC undergoing recombination with a compatible recombination site in the
 CC presence of one or more proteins active in lambda recombination. The
 CC nucleic acid of the invention replicates in prokaryotic and eukaryotic
 CC cells. The nucleic acid of the invention is useful for constructing a
 CC recombinant virus, generating replication-defective particles,
 CC preventing/inhibiting the expression of one or more genes in an organism,
 CC and are useful as gene therapy vectors. The nucleic acid of the invention
 CC is also useful for producing and expressing fusion polypeptides. The
 CC present DNA sequence represents a plasmid that was used in the
 CC exemplification of the invention.
 XX
 XX Sequence 5038 BP; 1342 A; 1178 C; 1198 G; 1320 T; 0 U; 0 Other;
 SQ
 Query Match 96.9%; Score 546.4; DB 12; Length 5038;
 Best Local Similarity 99.8%; Pred. No. 1.7e-166;
 Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CATGATGATAACAATGATGGTCTAATGTGCTTCAACCAACAATTCGTGTAACCTGTG 60
 DB 1 CATGATGATAACAATGATGGTCTAATGTGCTTCAACCAACAATTCGTGTAACCTGTG 60
 QY 61 TTTTTCATGTTGCCCAACAGCACCTTTTATCTCGGTGGCCCTCCACCACCACTTTT 120
 DB 61 TTTTTCATGTTGCCCAACAGCACCTTTTATCTCGGTGGCCCTCCACCACCACTTTT 120
 QY 121 GCATCTGCAAAAACACAGCTTTTGACGCGGGCCCATACATAGTACAACTCTAGTTTC 180
 DB 121 GCATCTGCAAAAACACAGCTTTTGACGCGGGCCCATACATAGTACAACTCTAGTTTC 180
 QY 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATAGCTTCCAAATACACTACACAC 240
 DB 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATAGCTTCCAAATACACTACACAC 240
 QY 241 ATTGAACCTTTTTCAGTGCAAAAAGTACGTGTGGCAGTACCGTAGCGCGGCTTATC 300
 DB 241 ATTGAACCTTTTTCAGTGCAAAAAGTACGTGTGGCAGTACCGTAGCGCGGCTTATC 300
 QY 301 GGGTCGGCTCTGTGTCAGTACGAATCATTATCGGACCGGACGAGTGTGTCTTATCGT 360
 DB 301 GGGTCGGCTCTGTGTCAGTACGAATCATTATCGGACCGGACGAGTGTGTCTTATCGT 360
 QY 361 GACAGGACCGCAGCTCTCTGTGTTGCTTAACCGGACCGGACGAGTCTCTTATCGGAACA 420
 DB 361 GACAGGACCGCAGCTCTCTGTGTTGCTTAACCGGACCGGACGAGTCTCTTATCGGAACA 420

QY 421 GGACGGCTCCATATCAGCGCGGTTTATCTCATGCGCGTGACCGGACACGAGGCGGCC 480
 DB 421 GGACGGCTCCATATCAGCGCGGTTTATCTCATGCGCGTGACCGGACACGAGGCGGCC 480
 QY 481 GTCCCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAG 540
 DB 481 GTCCCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAG 540
 QY 541 CATCTGTT 548
 DB 541 CATCTGTT 548
 RESULT 5
 AAT13730/C
 ID AAT13730 standard; DNA; 279 BP.
 XX
 AC AAT13730;
 DT 16-OCT-2003 (revised)
 DT 06-SEP-1996 (first entry)
 XX
 DE AcNPV ORF 152, residues 132387-132109.
 KW Autographa californica nuclear polyhedrosis virus clone 6; disruption;
 KW non-essential gene; heterologous protein production; expression vector;
 KW baculovirus; ss.
 OS Autographa californica nucleopolyhedrovirus; clone 6.
 PH Key Location/Qualifiers
 CDS 1..279
 FT /*tag= a
 FT /number= ORF 152
 FT /note= "corresponds to AcNPV nucleotides complement
 (132387-132109)"
 XX
 PN WO9601320-A2.
 XX
 PD 18-JAN-1996.
 XX
 PF 30-JUN-1995; 95WO-IB000578.
 XX
 PR 04-JUL-1994; 94GB-00013420.
 XX
 PA (NATU-) NATURAL ENVIRONMENT RES COUNCIL.
 XX
 PI Bishop D, Possee R, Ayres M;
 XX WPI; 1996-087670/09.
 XX
 PT Autographa californica nuclear polyhedrosis virus complete genome
 PT sequence - useful in the prodn. of vectors for enhanced heterologous
 PT protein expression, such as interleukin(s), interferon(s) and
 PT neurotoxin(s).
 XX
 PS Claim 1; Page 90-186; 122pp; English.
 XX
 CC AAT13636-731 show open reading frames 13, 20, 22-30, 32, 38, 41-46, 50-
 CC 60, 62-63, 66, 68-79, 81-87, 91-92, 96-98, 101-103, 106-127, 129-130, 140
 CC - 146, 148-150, 152 and 154 from a total of 154 ORFs identified in the
 CC Autographa californica nuclear polyhedrosis virus (AcNPV) clone 6. Each
 CC gene is numbered according to its position in the virus genome beginning
 CC at the left end of the linear map, and irrespective of its orientation.
 CC The direction of transcription is relative to that of the polyhedrin
 CC gene. Of the 154 ORFs identified it was found that some of the ORFs (ORFs
 CC 27, 30, 32, 71, 86, 123, 126 and 127) are dispensable for virus
 CC replication in cell culture or insect larvae. These genes can be deleted
 CC from the genome to: (a) provide additional sites for inserting single or
 CC multiple copies of foreign genes; and (b) to reduce the size of the virus
 CC complementary strand relative to the polyhedrin gene. The present
 CC sequence is designated ORF 152, and is on the complementary strand
 CC relative to the polyhedrin gene. (Updated on 16-OCT-2003 to standardise

CC OS field)

XX SQ Sequence 279 BP; 88 A; 36 C; 68 G; 87 T; 0 U; 0 Other;

Query Match 18.0%; Score 101.4; DB 2; Length 279;

Best Local Similarity 75.4%; Pred. No. 2.5e-22;

Matches 126; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 112 AACTTTTTGCACTGCAAAAAAACACGCTTTTGCACGGGGCCCATACATAGTACAAACT 171

Db 276 AATTTTTTGCATGCAAAAAAGTTCACCTTTTGCGCTGCACATCCCATATACAGTACAATCT 217

Qy 172 CTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCCAATACA 231

Db 216 CTCAAAATCGTAGACTATTTTATTAGAAATAGTCTACACTGTACGATACGCTCCCAATATA 157

Qy 232 CTACCACATTTGAACCTTTTTCAGTGCAGGCAAAAAAGTACGTCGCGC 278

Db 156 CTACTACATCACTTTTTCGATTTACAAAAAGTTCATTTTTCG 110

RESULT 6

AAT13635/c

ID AAT13635 standard; DNA; 133894 BP.

XX AC AAT13635;

XX AC AAT13635;

XX DT 16-OCT-2003 (revised)

XX DT 03-SEP-1996 (first entry)

XX DE AcNPV genomic DNA clone 6.

XX XX

XX Autographa californica nuclear polyhedrosis virus clone 6; disruption;

XX KW non-essential gene; heterologous protein production; expression vector;

XX KW baculovirus; ss.

XX XX

XX Autographa californica nucleopolyhedrovirus; clone 6.

XX OS

XX PN WO9601320-A2.

XX PD 18-JAN-1996.

XX PF 30-JUN-1995; 95WO-IB000578.

XX PR 04-JUL-1994; 94GB-00013420.

XX XX

PA (NATU-) NATURAL ENVIRONMENT RES COUNCIL.

XX Bishop D, Possee R, Ayres M;

XX PI WPI; 1996-087670/09.

XX DR

XX PT Autographa californica nuclear polyhedrosis virus complete genome

XX PT sequence - useful in the prodn. of vectors for enhanced heterologous

XX PT protein expression, such as interleukin(s), interferon(s) and

XX PT neurotoxin(s).

XX XX

XX Disclosure; Page 90-186; 122pp; English.

XX CC

XX The complete nucleotide sequence of the genome of clone 6 of the

XX CC baculovirus Autographa californica nuclear polyhedrosis virus (AcNPV) has

XX CC been determined. The sequence is taken from the Genbank record 122858.

XX CC The patent specification claims a polynucleotide selected from open

XX CC reading frames (ORFs 13, 20, 22-26, 28-30, 32, 38, 41-46, 50-60, 62-63,

XX CC 66, 68-79, 81-87, 91-92, 96-98, 101-103, 106-126, 129-130, 140-146, 148-

XX CC 150, 152 and 154 from a total of 154 ORFs identified by the patentees.

XX CC See T13636-731. Expression vectors contg. the complete genomic sequence

XX CC of AcNPV, with the exception that at least one non-essential ORF is

XX CC disrupted or replaced are useful for the synthesis of heterologous

XX CC proteins. (Updated on 16-OCT-2003 to standardise OS field)

XX SQ Sequence 133894 BP; 39195 A; 27151 C; 27347 G; 40201 T; 0 U; 0 Other;

Query Match 18.0%; Score 101.4; DB 2; Length 133894;

Best Local Similarity 75.4%; Pred. No. 4.4e-21;

Matches 126; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 112 AACTTTTTGCACTGCAAAAAAACACGCTTTTGCACGGGGCCCATACATAGTACAAACT 171

Db 133884 AATTTTTTGCATGCAAAAAAGTTCACCTTTTGCGCTGCACATCCCATATACAGTACAATCT 132326

Qy 172 CTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCCAATACA 231

Db 132324 CTACAAATCGTAGACTATTTTATTAGAAATAGTCTACACTGTACGATACGCTCCCAATATA 132365

Qy 232 CTACCACATTTGAACCTTTTTCAGTGCAGGCAAAAAAGTACGTCGCGC 278

Db 132264 CTACTACATCACTTTTTCGATTTACAAAAAGTTCATTTTTCG 132218

RESULT 7

ADCS1646/c

ID ADCS1646 standard; DNA; 28413 BP.

XX AC ADCS1646;

XX DT 18-DEC-2003 (first entry)

XX DE BmNPV genomic DNA nucleotides 100001-128413.

XX KW ds; silkworm; silkworm movement suppression.

XX XX

XX Bombyx mori nuclear polyhedrosis virus.

XX PN JP2003024062-A.

XX PD 28-JAN-2003.

XX PF 10-JUL-2001; 2001JP-00209305.

XX PR 10-JUL-2001; 2001JP-00209305.

XX PA (RIKA) RIKAGAKU KENKYUSHO.

XX XX

XX WPI; 2003-516415/49.

XX XX

XX Novel Bombyx mori polyhedrosis virus in which open reading frame gene is

XX PT inactivated, or its variant, useful for suppressing movement of Bombyx

XX PT mori.

XX PS Claim 5; SEQ ID NO 3; 53pp; Japanese.

XX CC The invention relates to a Bombyx mori (silkworm) polyhedrosis virus in

XX CC which open reading frame 8 (ORF 8) gene is inactivated. The inactivated

XX CC orf8 gene is useful for suppressing movement of silkworm and efficiently

XX CC prevents the movement of a silkworm from a chamber in which it is raised.

XX CC The present sequence is used in the exemplification of the invention.

XX XX

SQ Sequence 28413 BP; 8394 A; 5713 C; 5940 G; 8366 T; 0 U; 0 Other;

Query Match 16.6%; Score 93.4; DB 10; Length 28413;

Best Local Similarity 72.5%; Pred. No. 8.6e-19;

Matches 121; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 112 AACTTTTTGCACTGCAAAAAAACACGCTTTTGCACGGGGCCCATACATAGTACAAACT 171

Db 22231 AATTTTTTGCATACAAAAAGTTCGCTATGTTTGACATATATATACAGTACGACT 22172

Qy 172 CTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCCAATACA 231

Db 22171 CTACAAATCGTAGACTATTTTATTAGAAATAGTCTACACTGTACTACTACGCTCTCAATATA 22112

Qy 232 CTACCACATTTGAACCTTTTTCAGTGCAGGCAAAAAAGTACGTCGCGC 278

Db 22111 CTACTACATCACTTTTTCGATTTACAAAAAGTTCATTTTTCG 22065

```
RESULT 8
ADQ48576
ID ADQ48576 standard; DNA; 141 BP.
XX
AC ADQ48576;
XX
DT 09-SEP-2004 (first entry)
XX
DE Viral vector-related plasmid pTB/V5-His-DEST recombination region #1.
XX
KW viral vector; recombination site; recombinant virus;
KW replication-defective particle generation; gene expression inhibition;
KW gene therapy vector; ds; plasmid; recombination region.
XX
OS Unidentified.
XX
PN WO2004009768-A2.
XX
PD 29-JAN-2004.
XX
PF 18-JUL-2003; 2003WO-US022437.
XX
PR 18-JUL-2002; 2002US-0396335P.
PR 26-JUL-2002; 2002US-0398617P.
PR 19-NOV-2002; 2002US-0427231P.
PR 24-MAR-2003; 2003US-0456496P.
PR 03-JUN-2003; 2003US-0474940P.
XX
PA (INVI-) INVITROGEN CORP.
PA (BENN/) BENNETT R P.
PA (WELC/) WELCH P J.
PA (HARW/) HARWOOD S.
PA (MADD/) MADDEN K.
PA (FRIM/) FRIMPONG K.
PA (FRAN/) FRANK K E.
XX
PI Bennett RP, Welch PV, Harwood S, Madden K, Frimpong K, Franke KE;
XX
WPI; 2004-132944/13.
XX
PT Novel nucleic acid molecule comprising all or a portion of one or more
PT viral genome and further comprising two or more recombination sites that
PT do not substantially recombine with each other, useful as gene therapy.
XX
PS Disclosure; Fig 17; 555pp; English.
XX
CC The invention comprises a nucleic acid molecule consisting of all or a
CC portion of one or more viral/baculoviral genomes and further containing
CC at least two recombination sites that do not substantially recombine with
CC each other. One or more of the recombination sites is capable of
CC undergoing recombination with a compatible recombination site in the
CC presence of one or more proteins active in lambda recombination. The
CC nucleic acid of the invention replicates in prokaryotic and eukaryotic
CC cells. The nucleic acid of the invention is useful for constructing a
CC recombinant virus, generating replication-defective particles,
CC preventing/inhibiting the expression of one or more genes in an organism,
CC and are useful as gene therapy vectors. The nucleic acid of the invention
CC is also useful for producing and expressing fusion polypeptides. The
CC present DNA sequence represents a recombination region of a plasmid that
CC was used in the exemplification of the invention.
XX
SQ Sequence 141 BP; 43 A; 36 C; 27 G; 34 T; 0 U; 1 Other;
Query Match 11.0%; Score 62; DB 12; Length 141;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 487 CTTATCGCGCCTTATAAATACAGCCCGCAACGATCTGTGTAACACAGTTGAACAGCATCTG 546
DB 1 CTTATCGCGCCTTATAAATACAGCCCGCAACGATCTGTGTAACACAGTTGAACAGCATCTG 60
OY 547 TT 548
```

```
Db 61 TT 62
||
RESULT 9
ABZ10059/c
ID ABZ10059 standard; DNA; 6289 BP.
XX
AC ABZ10059;
XX
DT 16-JAN-2003 (first entry)
XX
DE Haematopoietic cell proliferation disorder related DNA sequence #199.
XX
KW Human; haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200277272-A2.
XX
PD 03-OCT-2002.
XX
PR 26-MAR-2002; 2002WO-EP003401.
PR 26-MAR-2001; 2001US-0278333P.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI Schwöpe I, Ziebarth H;
XX
WPI; 2003-018942/01.
XX
PT Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX
PS Claim 28; SEQ ID NO 199; 117pp; English.
XX
CC The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used; for
CC differentiating between healthy haematopoietic cells and proliferative
CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related
CC sequences. The nucleotide sequences from the present invention can also
CC be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferative disorders. The present method enables a
CC highly specific classification of haematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients
XX
SQ Sequence 6289 BP; 1680 A; 178 C; 1488 G; 2943 T; 0 U; 0 Other;
Query Match 7.0%; Score 39.6; DB 8; Length 6289;
Best Local Similarity 53.2%; Pred. No. 0.14;
Matches 84; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
OY 112 AACCTTTTTCACCTGCAAAAAAACACGCTTTTGACGCGGGCCCATACATAGTACAAACT 171
```

```

Db 4896 AATTTATTATATAACCAATAAACTATTATAAATCCACTACCAATATCATTTCAAAA 4837
QY 172 CTACGTTTCGTAGACTATTTCACATAAATAGTCTACACCGTTGTATACGTCCTCAATACA 231
Db 4836 TAATAATTTTAAACAATTATTATATAAATAAATCGTCACAATCTTTTACTCAAAAATA 4777
QY 232 CTACCAACACATTGAACCTTTTTCAGTCGCAAAAAGTA 269
Db 4776 ATACAACCTTTCCTCTCCCTATATAAAGCAAAAAAATA 4739

RESULT 10
ID AAS46501/c
ID AAS46501 standard; DNA; 9289 BP.
XX
AC AAS46501;
XX
DT 18-DEC-2001 (first entry)
XX
DE Tumour suppressor gene derived chemically modified sequence #223.
XX
KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;
KW tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.
XX
OS Homo sapiens.
XX
PN WO200168912-A2.
XX
PD 20-SEP-2001.
XX
PF 15-MAR-2001; 2001WO-EP002955.
XX
PR 15-MAR-2000; 2000DE-01013847.
PR 06-APR-2000; 2000DE-01019058.
PR 07-APR-2000; 2000DE-01019173.
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
WPI; 2001-602752/68.
XX
DR Fragments of chemically modified genes associated with tumor suppressor
PT genes and oncogenes, useful in designing primers and probes for analyzing
PT diseases associated with cytosine methylation state e.g. cancer.
XX
PS Claim 1; SEQ ID NO 223; 27pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and oncogenes
CC having a sequence taken from 536 (actually 533 since numbers 408, 458 and
CC 500 are missing from the sequence listing) sequences (Ss) and sequences
CC complementary to (Ss). The nucleic acid may be a peptide nucleic acid-
CC oligomer (PNA) of at least 9 nucleotides and may form part of a set of
CC probes for detecting the cytosine methylation state and/or single
CC nucleotide polymorphisms and also to be used in an array for analysing
CC diseases associated with CpG dinucleotides e.g. cancers and tumours. The
CC probes can also be used in a method for ascertaining genetic and/or
CC epigenetic parameters for the diagnosis and/or therapy of existing
CC diseases or the predisposition to specific diseases, by analysing
CC cytosine methylations. The parameters may be compared to another set of
CC genetic and/or epigenetic parameters, the differences serving as basis
CC for diagnosis and/or prognosis events which are disadvantageous to
CC patients. The present sequence is one of the 533 genomic sequences
CC derived from tumour suppressor genes and oncogenes. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ffp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 9289 BP; 2515 A; 206 C; 2097 G; 4471 T; 0 U; 0 Other;
Query Match 7.0%; Score 39.6; DB 4; Length 9289;
Best Local Similarity 53.2%; Pred. No. 0.16; Indels 0; Gaps 0;
Matches 84; Conservative 0; Mismatches 74;

QY 112 AACTTTTTTGCACACTGCAAAAAACACGCTTTTTCACGCGGCGCCATACATAGTACAAACT 171
Db 7896 AATTTATTATATAACCAATAAATACTATTATAAATCCACTACCAATATCATTTCAAAA 7837
QY 172 CTACGTTTCGTAGACTATTTCACATAAATAGTCTACACCGTTGTATACGTCCTCAATACA 231
Db 7836 TAATAATTTTAAACAATTATTATATAAATAAATCGTCACAATCTTTTACTCAAAAATA 7777
QY 232 CTACCAACACATTGAACCTTTTTCAGTCGCAAAAAGTA 269
Db 7776 ATACAACCTTTCCTCTCCCTATATAAAGCAAAAAAATA 7739

RESULT 11
ADE84121/c
ID ADE84121 standard; DNA; 9289 BP.
XX
AC ADE84121;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human lymphoid cell proliferative disorder gene derived DNA #57.
XX
KW ds; lymphoid cell proliferative disorder; methylation;
KW methylated CpG dinucleotide; single nucleotide polymorphism; SNP;
KW diffuse large B-cell lymphoma; mantle cell lymphoma;
KW chronic lymphocytic leukemia; small lymphocytic lymphoma;
KW follicular lymphoma; diagnosis; prognosis.
XX
OS Homo sapiens.
XX
PN WO2003044226-A2.
XX
PD 30-MAY-2003.
XX
PF 25-NOV-2002; 2002WO-EP013265.
XX
PR 23-NOV-2001; 2001DE-01057491.
PR 28-DEC-2001; 2001DE-01064501.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
PI Burger M, Caldwell C, Genc B, Becker E, Maier S, Nimmrich I;
XX
WPI; 2003-457621/43.
XX
DR Detecting and differentiating between lymphoid cell proliferative
PT disorders comprises contacting a target nucleic acid with at least one
PT reagent that distinguishes between methylated and non-methylated CpG
PT dinucleotides.
XX
PS Claim 26; SEQ ID NO 117; 448pp; English.
XX
CC The invention relates to a method of detecting and differentiating
CC between lymphoid cell proliferative disorders associated with at least
CC one gene and/or their regulatory regions in a subject by contacting a
CC target nucleic acid in a biological sample obtained from the subject with
CC at least one reagent or series of reagents that distinguish between
CC methylated and non-methylated CpG dinucleotides within the target nucleic
CC acid. The genes and/or their regulatory regions are preferably selected
CC from MDR1, CSNK2B, EGR4, AR, CDK4, RB2, CDC25A, GP1b beta, MYOD1, CDH3,
CC MYCL1, ELK1, ABL1, APC, BCL2, CDH1, CDKN1A, CDKN2a, CDKN2B, FOS,
CC GSK3P1, HIC-1, MGMT, MLH1, MOS, MYC, PTEN, RBL2, TGFBR2, TP73, CDKN1C,
CC GSK3beta, ESRI, APAF1, BAK1, BAX or HOKA5. Oligomers, peptide nucleic
CC acid (PNA)-oligomers and/or isolated nucleic acids based on the sequences
CC of the genes are useful for detecting the methylation state of all the
```

CC CpG dinucleotides within one or more the sequences, or their complements,
 CC for determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs), and for differentiating at least two of the medical
 CC conditions such as diffuse large B-cell lymphoma, mantle cell lymphoma,
 CC chronic lymphocytic leukemia, small lymphocytic lymphoma and follicular
 CC lymphoma. They are also useful for detecting of a predisposition to,
 CC differentiation between subclasses, diagnosis, prognosis, treating and/or
 CC monitoring of lymphoid cell proliferative disorder. This sequence
 CC represents a nucleic acid of a pretreated genomic DNA derived from the
 CC above mentioned genes.

XX SQ Sequence 9289 BP; 2515 A; 206 C; 2097 G; 4471 T; 0 U; 0 Other;

Query Match 7.0%; Score 39.6; DB 10; Length 9289;

Best Local Similarity 53.2%; Pred. No. 0.16; 74; Indels 0; Gaps 0;

Matches 84; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 112 AACTTTTTCAGTGCAGGCGGGCCCATACATAGTACAAACT 171

DB 7896 AATTATTATATACCAATAAATACTATTATAATCCACTACCAATATATCAAAAA 7837

QY 172 CTACGTTTCGTAGACTATTATACATAAATAGTCTACCGTTGTATAGCTCAATACA 231

DB 7836 TAATAATTTTAAACATTATTATATAAAAAATCGCACAAATCTTTTACTCAAAAAATA 7777

QY 232 CTACACACATTCGAACCTTTTTCAGTGCAGGCGGCAAAAAAGTA 269

DB 7776 ATACAACTTCTCTCCCTATATAAACAAGCAAAAAATA 7739

RESULT 12

ABZ10205/c

ID ABZ10205 standard; DNA; 6289 BP.

XX AC ABZ10205;

XX DT 16-JAN-2003 (first entry)

XX DE Haematopoietic cell proliferation disorder related DNA sequence #345.

XX KW Human; haematopoietic cell proliferation disorder; cytostatic;

XX KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;

XX KW cytosine methylation state; gene; ds.

XX OS Homo sapiens.

XX PN WO20027272-A2.

XX PD 03-OCT-2002.

XX PF 26-MAR-2002; 2002WO-EP003401.

XX PR 26-MAR-2001; 2001US-0278333P.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;

XX PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;

XX PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;

XX PI Schwöpe I, Ziebarth H;

XX DR WPI; 2003-018942/01.

XX PT Detecting and differentiating between hematopoietic cell proliferative

XX PT disorders, comprises contacting a target nucleic acid with a reagent that

XX PT distinguishes between methylated and non-methylated CpG dinucleotides.

XX PS Claim 28; SEQ ID NO 345; 117pp; English.

XX CC The present invention describes a method for detecting and

XX CC differentiating between haematopoietic cell proliferative disorders

XX CC associated with at least 1 gene and/or their regulatory regions in a

XX CC subject. The method comprises contacting a target nucleic acid in a

CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy haematopoietic cells and proliferative
 CC disorder haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related DNA
 CC sequences. The nucleotide sequences from the present invention can also
 CC be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC haematopoietic cell proliferative disorders. The present method enables a
 CC highly specific classification of haematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients

SQ Sequence 6289 BP; 1680 A; 0 C; 1488 G; 3121 T; 0 U; 0 Other;

Query Match 6.7%; Score 38; DB 8; Length 6289;

Best Local Similarity 52.5%; Pred. No. 0.45;

Matches 83; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 112 AACTTTTTCAGTGCAGGCGGGCCCATACATAGTACAAACT 171

DB 4896 AATTATTATATACCAATAAATACTATTATAATCCACTACCAATATATCAAAAA 4837

QY 172 CTACGTTTCGTAGACTATTATACATAAATAGTCTACACCGTTGTATAGCTCCAAATACA 231

DB 4836 TAATAATTTTAAACATTATTATATAAATCATCACAATCTTTTACTCAAAAAATA 4777

QY 232 CTACACACATTCGAACCTTTTTCAGTGCAGGCGCAAAAAAGTA 269

DB 4776 ATACAACTTCTCTCCCTATATAAACAAGCAAAAAATA 4739

RESULT 13

ADE84197/c

ID ADE84197 standard; DNA; 9289 BP.

XX AC ADE84197;

XX DT 29-JAN-2004 (first entry)

XX DE Human lymphoid cell proliferative disorder gene derived DNA #133.

XX KW ds; lymphoid cell proliferative disorder; methylation;

XX KW methylated CpG dinucleotide; single nucleotide polymorphism; SNP;

XX KW diffuse large B-cell lymphoma; mantle cell lymphoma;

XX KW chronic lymphocytic leukemia; small lymphocytic lymphoma;

XX KW follicular lymphoma; diagnosis; prognosis.

XX OS Homo sapiens.

XX PN WO2003044226-A2.

XX PD 30-MAY-2003.

XX PF 25-NOV-2002; 2002WO-EP013265.

XX PR 23-NOV-2001; 2001DE-01057491.

XX PR 28-DEC-2001; 2001DE-01064501.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Burger M, Caldwell C, Genc B, Becker E, Maier S, Nimmrich I;

XX PI WPI; 2003-457621/43.

XX DR Detecting and differentiating between lymphoid cell proliferative

XX PT disorders comprises contacting a target nucleic acid with at least one

XX WO200228891-A2.
PN
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-FR003061.
XX
PR 04-OCT-2000; 2000FR-00012697.
XX
XX (INSP) INST PASTEUR.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Kunst F, Glaser P;
XX
XX WPI; 2002-332479/37.
DR
XX
XX New genomic sequences from Listeria species, useful for detection,
PT treatment and prevention of infection, also related polypeptides,
PT antibodies and modulators.
XX
XX Claim 14; SEQ ID NO 1528; 180pp; French.
PS
XX
XX The present invention relates to nucleic acid sequences (ABQ67188-
CC ABQ71212) from Listeria sp. The sequences are useful as probes and
CC primers for identification and/or detection of Listeria (e.g. as
CC contaminants in foods, or mutational analysis) and for analysis of gene
CC expression. Proteins encoded by the nucleic acid sequences can be used to
CC screen for compounds that modulate gene expression, replication and
CC pathogenicity of Listeria (potential therapeutic agents), also for
CC treating infections by Listeria, and are useful as immunogens in anti-
CC Listeria vaccines. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. (Updated
CC on 29-AUG-2003 to standardise OS field)
XX
SQ Sequence 986 BP; 295 A; 199 C; 157 G; 328 T; 0 U; 7 Other;

Query Match
Best Local Similarity 6.3%; Score 35.4; DB 6; Length 986;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
Qy 155 CATACATAGTACAAACTCTAGCTTTCCGTAGACTATTTTACATAAATAGCTACACCGTTG 214
Db 318 CAGATTAGAAAACATTTTATCGTTGCCAGGAAAATTTAAATATCAAGAGCACCTCTGTTG 259
Qy 215 TATACGCTCCAAATACACTACACACATTGAACCTTTTTCGAGTGCAGAAAAGTACG 271
Db 258 CAGAACATTATATAAAGTTATCATCTCAATTTTCGGTCTGTAGAGGAAG 202

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 04:19:16 ; Search time 187.525 Seconds
(without alignments)
4921.273 Million cell updates/sec

Title: US-09-896-888A-1
Perfect score: 564
Sequence: 1 catgatgataacaatgtat.....tggtagcgacacacatg 564

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata1/ina/5A COMB.seq.*
2: /cgn2_6/prodata1/ina/5B COMB.seq.*
3: /cgn2_6/prodata1/ina/6A COMB.seq.*
4: /cgn2_6/prodata1/ina/6B COMB.seq.*
5: /cgn2_6/prodata1/ina/PCUS COMB.seq.*
6: /cgn2_6/prodata1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	36.4	6.5	505	US-09-621-976-15639	Sequence 15639, A
C 2	34.4	6.1	855	US-09-148-545-89	Sequence 89, Appl
C 3	33.8	6.0	3083	US-08-693-308-1	Sequence 1, Appl
C 4	33.2	5.9	902	US-09-270-767-25113	Sequence 25113, A
C 5	33.2	5.9	5003	US-09-270-767-9827	Sequence 9827, Ap
C 6	32.2	5.7	601	US-09-949-016-103819	Sequence 103819, A
C 7	32.2	5.7	601	US-09-949-016-103820	Sequence 103820, A
C 8	32.2	5.7	601	US-09-949-016-103821	Sequence 103821, A
C 9	32.2	5.7	601	US-09-949-016-103863	Sequence 103863, A
C 10	32.2	5.7	601	US-09-949-016-103864	Sequence 103864, A
C 11	32.2	5.7	601	US-09-949-016-103865	Sequence 103865, A
C 12	32.2	5.7	601	US-09-949-016-160230	Sequence 160230, A
C 13	32.2	5.7	601	US-09-949-016-160231	Sequence 160231, A
C 14	32.2	5.7	601	US-09-949-016-160232	Sequence 160232, A
C 15	32.2	5.7	601	US-09-949-016-160274	Sequence 160274, A
C 16	32.2	5.7	601	US-09-949-016-160275	Sequence 160275, A
C 17	32.2	5.7	601	US-09-949-016-160276	Sequence 160276, A
C 18	32.2	5.7	78649	US-09-949-016-14619	Sequence 14619, A
C 19	32.2	5.7	78649	US-09-949-016-14620	Sequence 14620, A
C 20	32.2	5.7	78649	US-09-949-016-16227	Sequence 16227, A
C 21	32.2	5.7	78649	US-09-949-016-16228	Sequence 16228, A
C 22	31.4	5.6	54452	US-09-949-016-12642	Sequence 12642, A
C 23	31.4	5.6	54452	US-09-949-016-13003	Sequence 13003, A
C 24	31.2	5.5	35675	US-09-949-016-13505	Sequence 13505, A
C 25	31	5.5	549	US-09-902-540-6433	Sequence 6433, Ap
C 26	31	5.5	601	US-09-949-016-86304	Sequence 86304, A
C 27	31	5.5	601	US-09-949-016-86305	Sequence 86305, A

ALIGNMENTS

RESULT 1

US-09-621-976-15639/c
; Sequence 15639, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15639
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15639

Query Match	6.5%;	Score 36.4;	DB 4;	Length 505;
Best Local Similarity	11.0%;	Pred. No. 0.044;		
Matches	40;	Conservative 159;	Mismatches 165;	Indels 0; Gaps 0;
QY	126	GCAGAAACACAGCCTTTTCACGCGGCCCATACATAGTACAACTCTACGTTTCGTAGA	185	
DB	364	RMAAGGSGYCGMTSYTSGSKMTGRKSMTRKRMMTYSGMWMTSYKCTKTGKYTGWKSK	305	
QY	186	CTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACACACATTGA	245	
DB	304	KTRWCTTSRKYMMWMSGKSWARSWMSWMSMACWMSASAYRARRSMYGARRSM	245	
QY	246	ACCTTTTTCAGTCGCAAAAAGTACGTCGCGAGTCACGTAGCGCGGCTTATCGGGTC	305	
DB	244	RAGAGWRRARRGKRGKSSMWSMWSMWSAGKARMCRRMWSMWSMWSMWSMWSMWSM	185	
QY	306	GGTCTCTGTACGTAGCAATCATTTATCGGACCGGACGAGTGTCTTATCGTGACAG	365	
DB	184	SCRGTCAKMWRYARVAKRYASSMGKYMGCWYAKCARMYGYRSTRSTGSRGMYR	125	
QY	366	GACGCCAGCTTCTGTGTTGTTAAACGCGAGCGGCAACTCTTATCGGAACAGACG	425	
DB	124	RMYYNWKYMMWSMWSMWSMWSMWSMWSMWSMWSMWSMWSMWSMWSMWSMWSMWSM	65	
QY	426	CGCTTCATATCAGCGCGCGTTCATCTCATGCGCGTACCGGACGAGCGCGCGTCCC	485	
DB	64	SYRWMWSKWCYSCGCGYCCSACRCYCWTRRMKWSYSRKSRMCCRYMSAYRYSK	5	
QY	486	GCTT 489		

Db 4 RTKT 1

RESULT 2

US-09-148-545-89
; Sequence 89, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PZ001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 05-Sep-1997
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 280
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 89
LENGTH: 855

Query Match 6.1%; Score 34.4; DB 4; Length 855;
Best Local Similarity 55.6%; Pred. No. 0.3;
Matches 65; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 439 GCGCGCGTTATCTCATGCGGTGACGACGAGGCGCGTCCCGTTATCGGCT 498
DB 12 GCGCACCCGAGCTCAGGTCGTCGCCACCCACCAAGTTCAGTGCAGCAGGCGT 71
QY 499 ATAAATACAGCCGCAACGATCTGGTAAACACAGTTGAACAGCATCTGTTACAGCGA 555
DB 72 ATGCGTGCCCTCACCTGCGGTGCGACAGGACTTGGACTGCGGATGGCAGCGA 128

RESULT 3
US-08-693-308-1/c
Sequence 1, Application US/08693308
Patent No. 6447996
GENERAL INFORMATION:
APPLICANT: HABERT-ORTOLI, Estelle
APPLICANT: AMIRANOFF, Brigitte
APPLICANT: LOUET, Isabelle
TITLE OF INVENTION: GALANIN RECEPTOR, NUCLEIC ACIDS,
TITLE OF INVENTION: TRANSFORMED CELLS AND USES THEREOF
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Rd. 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/693,308
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94/01808
FILING DATE: 17-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR95/00172
FILING DATE: 14-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D. Julie K.
REGISTRATION NUMBER: 38,619
REFERENCE/DOCKET NUMBER: ST94008-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610)454-3839
TELEFAX: (610)454-3808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3083 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 787..1836
US-08-693-308-1

Query Match 6.0%; Score 33.8; DB 3; Length 3083;
Best Local Similarity 54.4%; Pred. No. 1;
Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
QY 10 AAACAATGATGGTCTAATGTGCTTCAACAACAATCTGTGTGAACCTGTTTCATGT 69
DB 2939 AAATAATGTTGGTGGAAATGTTCAATTAACAATCATTACAAATTTTAAAGAAAGCATGC 2880
QY 70 TTGCCAACAGCACCTTTATCTCGTGGCTCCCAACCACTTTTTCGACTGCAA 129
DB 2879 ATTGTACAAGATAGGATACCCAGGTATATTTCAGTCCCATGATTTTTTGAATGTCAT 2820
QY 130 AAAA 134
DB 2819 CTTAA 2815

```
RESULT 4
US-09-270-767-25113/c
; Sequence 25113, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25113
; LENGTH: 902
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-25113

Query Match      5.9%; Score 33.2; DB 4; Length 902;
Best Local Similarity 56.4%; Pred. No. 0.79;
Matches 62; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 205 TACACCGTTGTATACGCTCCAAATACACTACACACATTGAACCTTTTTCAGTGCAGAAA 264
Db 440 TAACTCGATGTAGCATCTCCAGATACAATTGAGCTCATAGAAGCTGTTGAAAGGGAAT 381

QY 265 AAGTACGTGTGGCAGTACGTAGCGCGCCCTTATCGGTCGCGTCCTGT 314
Db 380 ATTCCGAATCTGTTGTGACAGCGCGCTTGTGAACCCGTCGTCGCTTT 331

RESULT 5
US-09-270-767-9827/c
; Sequence 9827, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9827
; LENGTH: 5003
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-9827

Query Match      5.9%; Score 33.2; DB 4; Length 5003;
Best Local Similarity 56.4%; Pred. No. 2.2;
Matches 62; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 205 TACACCGTTGTATACGCTCCAAATACACTACACACATTGAACCTTTTTCAGTGCAGAAA 264
Db 1212 TAACTCGATGTAGCATCTCCAGATACAATTGAGCTCATAGAAGCTGTTGAAAGGGAAT 1153

QY 265 AAGTACGTGTGGCAGTACGTAGCGCGCCCTTATCGGTCGCGTCCTGT 314
Db 1152 ATTCCGAATCTGTTGTGACAGCGCGCTTGTGAACCCGTCGTCGCTTT 1103

RESULT 6
US-09-949-016-103819
; Sequence 103819, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103819
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-103819

Query Match      5.7%; Score 32.2; DB 4; Length 601;
Best Local Similarity 56.0%; Pred. No. 1.4;
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 7 GATAAACAATGATGGTGCTTAATGCTTCTTCAACAACAATCTGTTGAACCTGTTTCA 66
Db 463 GCTAAACATAGATATTGAGACTTTTCTTCAGGAATTAATGATCTGAAGTTGTAATTA 522

QY 67 TGTITGCCAACAGCACCTTTATATCTCGTGGCCTCCCCACCACT 115
Db 523 TCTATCAGATATGCTCTTTTCTAATCAATGTTCCCACTCAGATT 571

RESULT 7
US-09-949-016-103820
; Sequence 103820, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103820
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-103820

Query Match      5.7%; Score 32.2; DB 4; Length 601;
Best Local Similarity 56.0%; Pred. No. 1.4;
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 7 GATAAACAATGATGGTGCTTAATGCTTCTTCAACAACAATCTGTTGAACCTGTTTCA 66
Db 133 GCTAAACATAGATATTGAGACTTTTCTTCAGGAATTAATGATCTGAAGTTGTAATTA 192

QY 67 TGTITGCCAACAGCACCTTTATATCTCGTGGCCTCCCCACCACT 115
Db 193 TCTATCAGATATGCTCTTTTCTAATCAATGTTCCCACTCAGATT 241

RESULT 8
US-09-949-016-103821
; Sequence 103821, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,498
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 103821
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-103821

Query Match 5.7%; Score 32.2; DB 4; Length 601;
Best Local Similarity 56.0%; Pred. No. 1.4;
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
QY 7 GATAAACAATGTATGGTCTTAATGTTGCTTCAACAACAATCTGTTGAACCTGTTTCA 66
DB 102 GCTAAACATAGATATTGAGACTTTTCTTCAGGAATTATTGATCTGAAGTTGATTAA 161
QY 67 TGTTCGCAACAGCACCCTTTATCTCGGTGGCTCCCAACCACT 115
DB 162 TCTATCAGATATGCTCTTTCTTAATCCATGTTGTTCCCATCTCAGATT 210

RESULT 9
US-09-949-016-103863
Sequence 103863, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 103863
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-103863

Query Match 5.7%; Score 32.2; DB 4; Length 601;
Best Local Similarity 56.0%; Pred. No. 1.4;
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
QY 7 GATAAACAATGTATGGTCTTAATGTTGCTTCAACAACAATCTGTTGAACCTGTTTCA 66
DB 463 GCTAAACATAGATATTGAGACTTTTCTTCAGGAATTATTGATCTGAAGTTGATTAA 522
QY 67 TGTTCGCAACAGCACCCTTTATCTCGGTGGCTCCCAACCACT 115
DB 523 TCTATCAGATATGCTCTTTCTTAATCCATGTTGTTCCCATCTCAGATT 571

RESULT 10
US-09-949-016-103864
Sequence 103864, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 103864
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-103864

Query Match 5.7%; Score 32.2; DB 4; Length 601;
Best Local Similarity 56.0%; Pred. No. 1.4;
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
QY 7 GATAAACAATGTATGGTCTTAATGTTGCTTCAACAACAATCTGTTGAACCTGTTTCA 66
DB 133 GCTAAACATAGATATTGAGACTTTTCTTCAGGAATTATTGATCTGAAGTTGATTAA 192
QY 67 TGTTCGCAACAGCACCCTTTATCTCGGTGGCTCCCAACCACT 115
DB 193 TCTATCAGATATGCTCTTTCTTAATCCATGTTGTTCCCATCTCAGATT 241

RESULT 11
US-09-949-016-103865
Sequence 103865, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 103865
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-103865

Query Match 5.7%; Score 32.2; DB 4; Length 601;
Best Local Similarity 56.0%; Pred. No. 1.4;
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
QY 7 GATAAACAATGTATGGTCTTAATGTTGCTTCAACAACAATCTGTTGAACCTGTTTCA 66
DB 102 GCTAAACATAGATATTGAGACTTTTCTTCAGGAATTATTGATCTGAAGTTGATTAA 161
QY 67 TGTTCGCAACAGCACCCTTTATCTCGGTGGCTCCCAACCACT 115
DB 162 TCTATCAGATATGCTCTTTCTTAATCCATGTTGTTCCCATCTCAGATT 210

RESULT 12
US-09-949-016-160230

```
; Sequence 160230, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 160230
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-160230

Query Match          5.7%; Score 32.2; DB 4; Length 601;
Best Local Similarity 56.0%; Pred. No. 1.4;
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 7 GATAAACAATGATGGTCTAATGTTGCTTCAACAACAATTCGTGTAACCTGTTTCA 66
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 463 GCTAAACATAGATATTGAGACTTTTCTTCAGGAATTTGATCTGAAGTTGTTAA 522

QY 67 TGTTCGCAACAAGCACCCTTTATCTCGTGGCCCTCCCAACCAACT 115
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 523 TCTATCAGAAATGCTCTTTCTTAATCAATGGTTCCTCATCTCAGATT 571

RESULT 13
US-09-949-016-160231
; Sequence 160231, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 160231
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-160231

Query Match          5.7%; Score 32.2; DB 4; Length 601;
Best Local Similarity 56.0%; Pred. No. 1.4;
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 7 GATAAACAATGATGGTCTAATGTTGCTTCAACAACAATTCGTGTAACCTGTTTCA 66
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 133 GCTAAACATAGATATTGAGACTTTTCTTCAGGAATTTGATCTGAAGTTGTTAA 192

QY 67 TGTTCGCAACAAGCACCCTTTATCTCGTGGCCCTCCCAACCAACT 115
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 193 TCTATCAGAAATGCTCTTTCTTAATCAATGGTTCCTCATCTCAGATT 241

RESULT 14
US-09-949-016-160232
; Sequence 160232, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 160232
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-160232

Query Match          5.7%; Score 32.2; DB 4; Length 601;
Best Local Similarity 56.0%; Pred. No. 1.4;
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 7 GATAAACAATGATGGTCTAATGTTGCTTCAACAACAATTCGTGTAACCTGTTTCA 66
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102 GCTAAACATAGATATTGAGACTTTTCTTCAGGAATTTGATCTGAAGTTGTTAA 161

QY 67 TGTTCGCAACAAGCACCCTTTATCTCGTGGCCCTCCCAACCAACT 115
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 162 TCTATCAGAAATGCTCTTTCTTAATCAATGGTTCCTCATCTCAGATT 210

RESULT 15
US-09-949-016-160274
; Sequence 160274, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 160274
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-160274

Query Match          5.7%; Score 32.2; DB 4; Length 601;
Best Local Similarity 56.0%; Pred. No. 1.4;
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 7 GATAAACAATGATGGTCTAATGTTGCTTCAACAACAATTCGTGTAACCTGTTTCA 66
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 463 GCTAAACATAGATATTGAGACTTTTCTTCAGGAATTTGATCTGAAGTTGTTAA 522

QY 67 TGTTCGCAACAAGCACCCTTTATCTCGTGGCCCTCCCAACCAACT 115
```

Db 523 TCTATCAGAAATATGCTCTTTTCTAAATCCCATTTGGTTCCCCCATCTCAGATT 571

Search completed: May 10, 2005, 07:29:33
Job time : 189.525 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 05:52:55 ; Search time 1593.6 Seconds
(without alignments)
2164.037 Million cell updates/sec

Title: US-09-896-888A-1

Perfect score: 564

Sequence: 1 catgatgataaacaatgtat.....tggtacagcgacacaacatg 564

Scoring table: IDENTITY_NVC

Gapop 10.0 , Gapext 1.0

Searched: 5654200 seqs, 3057283753 residues

Total number of hits satisfying chosen parameters: 11308400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	564	100.0	564	9	US-09-896-888A-1
2	548	97.2	560	18	US-10-622-088-126
3	548	97.2	2773	16	US-10-295-074-60
4	548	97.2	2773	18	US-10-846-911-60
5	546.4	96.9	5038	18	US-10-622-088-89
6	462	81.9	462	9	US-09-896-888A-14
7	62	11.0	147	18	US-10-622-088-127
8	62	11.0	325	18	US-10-622-088-149
9	57	10.1	92	9	US-09-896-888A-16
10	42.6	7.6	88	9	US-09-896-888A-15
11	39.6	7.0	6289	18	US-10-473-126-199

c 12	39.6	7.0	9289	17	US-10-221-714A-223	Sequence 223, App
c 13	38	6.7	6289	18	US-10-473-126-345	Sequence 345, App
c 14	35.4	6.3	986	17	US-10-398-221-1528	Sequence 1528, Ap
c 15	35.4	6.3	1549	17	US-10-398-221-3152	Sequence 3152, Ap
c 16	34.4	6.1	855	9	US-09-981-876-89	Sequence 89, Appl
c 17	34.4	6.1	855	10	US-09-148-545-89	Sequence 89, Appl
c 18	34.2	6.1	436	18	US-10-425-115-104158	Sequence 104158,
c 19	33.8	6.0	2117	13	US-10-087-192-1463	Sequence 1463, Ap
c 20	33.8	6.0	3056	15	US-10-225-567A-125	Sequence 125, App
c 21	33.8	6.0	3083	16	US-10-166-568-1	Sequence 1, Appli
c 22	33.8	6.0	39344	13	US-10-087-192-1462	Sequence 1462, Ap
c 23	33.2	5.9	621	18	US-10-437-963-77138	Sequence 77138, A
c 24	33	5.9	14429	15	US-10-311-455-2215	Sequence 2215, Ap
c 25	33	5.9	14429	18	US-10-433-793-127	Sequence 127, App
c 26	32.8	5.8	454	9	US-09-770-441-580	Sequence 580, App
c 27	32.8	5.8	2000	9	US-09-938-842A-5148	Sequence 5148, Ap
c 28	32.8	5.8	2000	11	US-09-938-842A-5148	Sequence 5148, Ap
c 29	32.6	5.8	469	18	US-10-425-115-102790	Sequence 102790,
c 30	32.6	5.8	495	10	US-09-814-353-5212	Sequence 5212, Ap
c 31	32.6	5.8	495	10	US-09-814-353-11499	Sequence 11499, A
c 32	32.6	5.8	60604	17	US-10-300-263-11	Sequence 11, Appli
c 33	32.4	5.7	137870	17	US-10-351-951-1	Sequence 1, Appli
c 34	32.2	5.7	802	14	US-10-184-644-312	Sequence 312, App
c 35	32.2	5.7	802	14	US-10-184-634-312	Sequence 312, App
c 36	32.2	5.7	965	17	US-10-369-493-30240	Sequence 30240, A
c 37	32.2	5.7	3138	15	US-10-156-761-2538	Sequence 2538, Ap
c 38	32.2	5.7	143412	13	US-10-087-192-997	Sequence 997, App
c 39	32.2	5.7	3673778	16	US-10-312-841-1	Sequence 1, Appli
c 40	32.2	5.7	9025608	15	US-10-156-761-1	Sequence 1, Appli
c 41	32	5.7	1875	14	US-10-198-846-11349	Sequence 11349, A
c 42	32	5.7	3297	17	US-10-369-493-30372	Sequence 30372, A
c 43	31.8	5.6	180	17	US-10-424-599-38274	Sequence 38274, A
c 44	31.8	5.6	441	18	US-10-425-115-55004	Sequence 55004, A
c 45	31.8	5.6	592	17	US-10-424-599-101900	Sequence 101900,

ALIGNMENTS

RESULT 1

US-09-896-888A-1
; Sequence 1, Application US/09896888A
; Patent No. US20020116723A1
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/896,888A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/048,911
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/049,946
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Orgyia pseudotsugata
US-09-896-888A-1

Query Match 100.0%; Score 564; DB 9; Length 564;
Best Local Similarity 100.0%; Pred. No. 2.4e-175;
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGATGATAAACAATGATGCTGCTTCAACAACAATCTGTTGAACGTG 60

Db 1 CATGATGATAAACAATGATGCTGCTTCAACAACAATCTGTTGAACGTG 60

QY 61 TTTTCATGTTTCCCAACAGCACCTTATACTCGGTGGCTCCCAACCACTTTT 120

Db 61 TTTTCATGTTTCCCAACAGCACCTTATACTCGGTGGCTCCCAACCACTTTT 120

```
QY 121 GCATGCTCAAAAAACACGCTTTTTCACGCGCGGCCCATACATAGTACAACTCTACGTTTC 180
DB 121 GCATGCTCAAAAAACACGCTTTTTCACGCGCGGCCCATACATAGTACAACTCTACGTTTC 180
QY 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTTCCAAATACACTACCAAC 240
DB 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTTCCAAATACACTACCAAC 240
QY 241 ATTGAACCTTTTTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 300
DB 241 ATTGAACCTTTTTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 300
QY 301 GGGTCGCGCTTCTGTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 360
DB 301 GGGTCGCGCTTCTGTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 360
QY 361 GACAGCGCCAGCTTCTGTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 420
DB 361 GACAGCGCCAGCTTCTGTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 420
QY 421 GACAGCGCCAGCTTCTGTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 480
DB 421 GACAGCGCCAGCTTCTGTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 480
QY 481 GTCCCGCTTATCGCGCTTATTAATACAGCGCGGCTTATTAATACAGCGCGGCTTATTAATACAG 540
DB 481 GTCCCGCTTATCGCGCTTATTAATACAGCGCGGCTTATTAATACAGCGCGGCTTATTAATACAG 540
QY 541 CATCTGTTACAGCGACACAAATG 564
DB 541 CATCTGTTACAGCGACACAAATG 564
```

RESULT 2

```
US-10-622-088-126
; Sequence 126, Application US/10622088
; Publication No. US20040219516A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 126
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: OpIE2 promoter sequence
US-10-622-088-126
```

```
Query Match          97.2%; Score 548; DB 18; Length 560;
Best Local Similarity 100.0%; Pred. No. 4.7e-170;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CATGATGATAAACAAATGATGCTGCTATGTTGTTTCAACAACAATCTCTGTAACCTGTG 60
DB 5 CATGATGATAAACAAATGATGCTGCTATGTTGTTTCAACAACAATCTCTGTAACCTGTG 64
QY 61 TTTTCATGTTTGGCCAAACAGCACCTTTTATATCTCGGTGGCCTCCCCACCACTTTTTC 120
DB 65 TTTTCATGTTTGGCCAAACAGCACCTTTTATATCTCGGTGGCCTCCCCACCACTTTTTC 124
QY 121 GCATGCAAAAAAACAACGCTTTTTCACGCGCGGCCCATACATAGTACAACTCTACGTTTC 180
DB 125 GCATGCAAAAAAACAACGCTTTTTCACGCGCGGCCCATACATAGTACAACTCTACGTTTC 184
QY 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTTCCAAATACACTACCAAC 240
DB 185 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTTCCAAATACACTACCAAC 244
QY 241 ATTGAACCTTTTTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 300
DB 245 ATTGAACCTTTTTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 304
QY 301 GGGTCGCGCTTCTGTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 360
DB 305 GGGTCGCGCTTCTGTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 364
QY 361 GACAGCGCCAGCTTCTGTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 420
DB 365 GACAGCGCCAGCTTCTGTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 424
QY 421 GACAGCGCCAGCTTCTGTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 480
DB 425 GACAGCGCCAGCTTCTGTCAGTGCAGTCAAAAAAGTACGTCGCGAGTCACTAGCGCGGCTTATC 484
QY 481 GTCCCGCTTATCGCGCTTATTAATACAGCGCGGCTTATTAATACAGCGCGGCTTATTAATACAG 540
DB 485 GTCCCGCTTATCGCGCTTATTAATACAGCGCGGCTTATTAATACAGCGCGGCTTATTAATACAG 544
QY 541 CATCTGTT 548
DB 545 CATCTGTT 552
```

RESULT 3

US-10-295-074-60

; Sequence 60, Application US/10295074

; Publication No. US20030185845A1

; GENERAL INFORMATION:

; APPLICANT: Pharmexa A/S

; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS

; FILE REFERENCE: P1013DK00

; CURRENT APPLICATION NUMBER: US/10/295,074

; CURRENT FILING DATE: 2002-11-15

; NUMBER OF SEQ ID NOS: 60

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 60

; LENGTH: 2773

; TYPE: DNA

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: p22Op2F expression vector for insect cells

; NAME/KEY: misc recomb

; LOCATION: (561)..(566)

; OTHER INFORMATION: HindIII site

; FEATURE:

; NAME/KEY: misc recomb

; LOCATION: (573)..(578)

; OTHER INFORMATION: Aval site

; FEATURE:

; NAME/KEY: misc recomb

; LOCATION: (586)..(591)

; OTHER INFORMATION: EcoRI site

; FEATURE:

```

; NAME/KEY: misc_recomb
; LOCATION: (593)..(598)
; OTHER INFORMATION: BamHI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (625)..(630)
; OTHER INFORMATION: ClaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (629)..(634)
; OTHER INFORMATION: ClaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (1156)..(1161)
; OTHER INFORMATION: ApaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2128)..(2133)
; OTHER INFORMATION: PstI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2204)..(2209)
; OTHER INFORMATION: NcoI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2284)..(2289)
; OTHER INFORMATION: AvaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2294)..(2299)
; OTHER INFORMATION: AvaI, SmaI, and XmaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2551)..(2556)
; OTHER INFORMATION: ApaI site
;
US-10-295-074-60

Query Match          97.2%; Score 548; DB 16; Length 2773;
Best Local Similarity 100.0%; Pred. No. 1e-169;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGATGATTAACAAGTATGCTGCTTAATGCTGCTCAACAAACAATTCGTTGAACGTG 60
DB 5 CATGATGATTAACAAGTATGCTGCTTAATGCTGCTCAACAAACAATTCGTTGAACGTG 64
QY 61 TTTTCATGTTTGCACACGACCTTATATCTCGTGGCTCCCGACCACTTTT 120
DB 65 TTTTCATGTTTGCACACGACCTTATATCTCGTGGCTCCCGACCACTTTT 124
QY 121 GCACGTCAAAAAACACGCTTTTGCACGGGGCCATACATAGTACAAACTCTACGTTTC 180
DB 125 GCACGTCAAAAAACACGCTTTTGCACGGGGCCATACATAGTACAAACTCTACGTTTC 184
QY 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATAGCTCCAAATACACTACACAC 240
DB 185 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATAGCTCCAAATACACTACACAC 244
QY 241 ATTGAACCTTTTTCAGTGCACAAAGTAGTCTCGGCAGTACAGTAGCGCGCCTTATC 300
DB 245 ATTGAACCTTTTTCAGTGCACAAAGTAGTCTCGGCAGTACAGTAGCGCGCCTTATC 304
QY 301 GGGTCGGCTCTGTTCAGTACGAATCACATTATCGGACGGAGGTGTCTTATCGT 360
DB 305 GGGTCGGCTCTGTTCAGTACGAATCACATTATCGGACGGAGGTGTCTTATCGT 364
QY 361 GACAGGACGCGAGTCTCTGTGTGCTTAACCGGACGGAGCGCAACTCTTATCGGAACA 420
DB 365 GACAGGACGCGAGTCTCTGTGTGCTTAACCGGACGGAGCGCAACTCTTATCGGAACA 424
QY 421 GGAGCGCCTCCATATACGCGCGCGGTATCTCATGCGGTGACCGGACGAGGCGCC 480
DB 425 GGAGCGCCTCCATATACGCGCGCGGTATCTCATGCGGTGACCGGACGAGGCGCC 484

481 GTCCCGCTTATCGCGCTATAAATAACAGCCCGCAACGATCTGTTAAACACAGTTGAACAG 540
485 GTCCCGCTTATCGCGCTATAAATAACAGCCCGCAACGATCTGTTAAACACAGTTGAACAG 544
541 CATCTGTT 548
545 CATCTGTT 552

RESULT 4
US-10-846-911-60
; Sequence 60, Application US/10846911
; Publication No. US20040258660A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen
; APPLICANT: NIELSEN, Finn Stausholm
; APPLICANT: BRAV, Tomas
; APPLICANT: VOLDHORG, Bjorn
; APPLICANT: MOURITSEN, Soren
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: 674542-2018
; CURRENT APPLICATION NUMBER: US/10/846,911
; CURRENT FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: PCT/DK02/00764
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: 60/331,575
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: PA 2001 01702
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 2773
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p22Op2F expression vector for insect cells
; NAME/KEY: misc_recomb
; LOCATION: (561)..(566)
; OTHER INFORMATION: HindIII site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (573)..(578)
; OTHER INFORMATION: AvaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (586)..(591)
; OTHER INFORMATION: EcoRI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (593)..(598)
; OTHER INFORMATION: BamHI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (625)..(630)
; OTHER INFORMATION: ClaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (629)..(634)
; OTHER INFORMATION: ClaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (1156)..(1161)
; OTHER INFORMATION: ApaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2128)..(2133)
; OTHER INFORMATION: PstI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2204)..(2209)
; OTHER INFORMATION: NcoI site
```

```
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2284)..(2289)
; OTHER INFORMATION: Aval site
;
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2294)..(2299)
; OTHER INFORMATION: Aval, SmaI, and XmaI site
;
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2551)..(2556)
; OTHER INFORMATION: ApaLI site
US-10-846-911-60

Query Match      97.2%; Score 548; DB 18; Length 2773;
Best Local Similarity 100.0%; Pred. No. 1e-169;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGATGATAAACAATGTATGCTAATGTGCTTAATGTGCTTCAACAACAATTCGTGTAACCTGTG 60
Db 5 CATGATGATAAACAATGTATGCTAATGTGCTTAATGTGCTTCAACAACAATTCGTGTAACCTGTG 64

Qy 61 TTTTCATGTTTCCCAACAAGCACTTTTATCTCGTGGGCTCCCAACCAACTTTTTT 120
Db 65 TTTTCATGTTTCCCAACAAGCACTTTTATCTCGTGGGCTCCCAACCAACTTTTTT 124

Qy 121 GCATGCAAAAAAACAAGCACTTTTCAACGGGGCCCATACATAGTACAACTCTAGTTTC 180
Db 125 GCATGCAAAAAAACAAGCACTTTTCAACGGGGCCCATACATAGTACAACTCTAGTTTC 184

Qy 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCACAC 240
Db 185 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCACAC 244

Qy 241 ATTGAACCTTTTTCAGTGCAAAAAAGTACGTGCGGAGTCACTAGCGCGGCTTATC 300
Db 245 ATTGAACCTTTTTCAGTGCAAAAAAGTACGTGCGGAGTCACTAGCGCGGCTTATC 304

Qy 301 GGGTCGGCTCTGTACAGTACGAATCACAATTCAGGACGAGTCTGTCTATCGT 360
Db 305 GGGTCGGCTCTGTACAGTACGAATCACAATTCAGGACGAGTCTGTCTATCGT 364

Qy 361 GACAGCGCCAGCTCTCTGTGTTGCTTAACCGCAGCGGAGCACTCTTATCGGAACA 420
Db 365 GACAGCGCCAGCTCTCTGTGTTGCTTAACCGCAGCGGAGCACTCTTATCGGAACA 424

Qy 421 GACGCGCTCCATATCAGCGCGGCTTATCTCATGCGGTGACCGGACAGAGCGGCC 480
Db 425 GACGCGCTCCATATCAGCGCGGCTTATCTCATGCGGTGACCGGACAGAGCGGCC 484

Qy 481 GTCCCGCTTATCGCGCTTATAAATACAGCGCGGAGCACTGTTAAACACAGTTGAACAG 540
Db 485 GTCCCGCTTATCGCGCTTATAAATACAGCGCGGAGCACTGTTAAACACAGTTGAACAG 544

Qy 541 CATCTGTT 548
Db 545 CATCTGTT 552
```

RESULT 5

```
US-10-622-088-89
; Sequence 89, Application US/10622088
; Publication No. US20040219516A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth E.
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
```

```
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 89
; LENGTH: 5038
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PIB/V5-His-DEST
US-10-622-088-89
```

Query Match 96.9%; Score 546.4; DB 18; Length 5038;

Best Local Similarity 99.8%; Pred. No. 4.7e-169;

Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 CATGATGATAAACAATGTATGCTAATGTGCTTCAACAACAATTCGTGTAACCTGTG 60
Db 1 CATGATGATAAACAATGTATGCTAATGTGCTTCAACAACAATTCGTGTAACCTGTG 60

Qy 61 TTTTCATGTTTCCCAACAAGCACTTTTATCTCGTGGGCTCCCAACCAACTTTTTT 120
Db 61 TTTTCATGTTTCCCAACAAGCACTTTTATCTCGTGGGCTCCCAACCAACTTTTTT 120

Qy 121 GCATGCAAAAAAACAAGCACTTTTCAACGGGGCCCATACATAGTACAACTCTAGTTTC 180
Db 121 GCATGCAAAAAAACAAGCACTTTTCAACGGGGCCCATACATAGTACAACTCTAGTTTC 180

Qy 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCACAC 240
Db 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCACAC 240

Qy 241 ATTGAACCTTTTTCAGTGCAAAAAAGTACGTGCGGAGTCACTAGCGCGGCTTATC 300
Db 241 ATTGAACCTTTTTCAGTGCAAAAAAGTACGTGCGGAGTCACTAGCGCGGCTTATC 300

Qy 301 GGGTCGGCTCTGTACAGTACGAATCACAATTCAGGACGAGTCTGTCTATCGT 360
Db 301 GGGTCGGCTCTGTACAGTACGAATCACAATTCAGGACGAGTCTGTCTATCGT 360

Qy 361 GACAGCGCCAGCTCTCTGTGTTGCTTAACCGGACGCGGAGCACTCTTATCGGAACA 420
Db 361 GACAGCGCCAGCTCTCTGTGTTGCTTAACCGGACGCGGAGCACTCTTATCGGAACA 420

Qy 421 GACGCGCTCCATATCAGCGCGGCTTATCTCATGCGGTGACCGGACAGAGCGGCC 480
Db 421 GACGCGCTCCATATCAGCGCGGCTTATCTCATGCGGTGACCGGACAGAGCGGCC 480

Qy 481 GTCCCGCTTATCGCGCTTATAAATACAGCGCGGAGCACTGTTAAACACAGTTGAACAG 540
Db 481 GTCCCGCTTATCGCGCTTATAAATACAGCGCGGAGCACTGTTAAACACAGTTGAACAG 540

Qy 541 CATCTGTT 548
Db 541 CATCTGTT 548
```

RESULT 6

```
US-09-896-888A-14
; Sequence 14, Application US/09896888A
; Patent No. US20020116723A1
; GENERAL INFORMATION:
```

```
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/896,888A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/048,911
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/049,946
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Promoter
; OTHER INFORMATION: sequence of the OpNPV ie2 gene
US-09-896-888A-14

Query Match      81.9%; Score 462; DB 9; Length 462;
Best Local Similarity 100.0%; Pred. No. 1.1e-141;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 CCCACACCAACTTTTTCATCTGCAAAACACGCTTTTCACGCGGGCCATACAT 161
DB 1 CCCACACCAACTTTTTCATCTGCAAAACACGCTTTTCACGCGGGCCATACAT 60

QY 162 AGTACAACTCTACGTTTGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGC 221
DB 61 AGTACAACTCTACGTTTGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGC 120

QY 222 TCCAAATACATACACACATTTGAACCTTTTTCAGTGCAGAAAGTACGTGTGCGCAGT 281
DB 121 TCCAAATACATACACACATTTGAACCTTTTTCAGTGCAGAAAGTACGTGTGCGCAGT 180

QY 282 CAGTAGGCGGCGCTTATCGGTGCGGTCTCTGTCAGTAGGAATACATTTTCGACCGG 341
DB 181 CAGTAGGCGGCGCTTATCGGTGCGGTCTCTGTCAGTAGGAATACATTTTCGACCGG 240

QY 342 ACAGAGTGTCTTATCGTGACAGGACGCGCTTCTGTTGTTGCTAAACGACGCGGAC 401
DB 241 ACAGAGTGTCTTATCGTGACAGGACGCGCTTCTGTTGTTGCTAAACGACGCGGAC 300

QY 402 GCAACTCCTTATCGGAACGAGCGCGCTCCATATCAGCGCGCTTATCTCATGCGCGT 461
DB 301 GCAACTCCTTATCGGAACGAGCGCGCTCCATATCAGCGCGCTTATCTCATGCGCGT 360

QY 462 GACCGACACGAGCGCGCGCTCCGCTTATCGCGCTTATAAATACAGCGCGCAACGATCT 521
DB 361 GACCGACACGAGCGCGCGCTCCGCTTATCGCGCTTATAAATACAGCGCGCAACGATCT 420

QY 522 GGTAAACAGTTGACAGCATCTGTTACAGCAGACACAT 563
DB 421 GGTAAACAGTTGACAGCATCTGTTACAGCAGACACAT 462

RESULT 7
US-10-622-088-127
; Sequence 127, Application US/10622088
; Publication No. US20040219516A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 149
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Recombination region of pIB/V5 His DEST
```

```
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 127
; LENGTH: 147
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombination region of pIB/V5 His-DEST
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (141)..(148)
; OTHER INFORMATION: n may be any nucleotide
US-10-622-088-127

Query Match      11.0%; Score 62; DB 18; Length 147;
Best Local Similarity 100.0%; Pred. No. 8.2e-10;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 487 CTTATCGCGCTTAAATACAGCCGCAACGATCTGTTAAACACAGTTGAACAGCATCTG 546
DB 1 CTTATCGCGCTTAAATACAGCCGCAACGATCTGTTAAACACAGTTGAACAGCATCTG 60

QY 547 TT 548
DB 61 TT 62

RESULT 8
US-10-622-088-149
; Sequence 149, Application US/10622088
; Publication No. US20040219516A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth E.
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 149
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Recombination region of pIB/V5 His DEST
```

FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (141)..(142)
; OTHER INFORMATION: n is a, c, g, or t
FEATURE:
; NAME/KEY: CDS
; LOCATION: (145)..(276)
US-10-622-088-149

Query Match 11.0%; Score 62; DB 18; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 487 CTTATCGCGCTATAAATACAGCCGCAACGATCTGGTAACACACAGTGAACAGCATCTG 546
|||
Db 1 CTTATCGCGCTATAAATACAGCCGCAACGATCTGGTAACACACAGTGAACAGCATCTG 60
QY 547 TT 548
|||
Db 61 TT 62

RESULT 9

US-09-896-888A-16
; Sequence 16, Application US/09896888A
; Patent No. US20020116723A1
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/896,888A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/048,911
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/049,946
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 92
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fragment of
; OTHER INFORMATION: promoter sequence of the AcMNPV ien gene
US-09-896-888A-16

Query Match 10.1%; Score 57; DB 9; Length 92;
Best Local Similarity 77.5%; Pred. No. 2.9e-08;
Matches 69; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 186 CTTATTTACATAAATAGTCTACACCGTTGTATACGCTCAATACACTACACACATGCA 245
|||
Db 1 CTTATTTATTAAGATGCTACACTGTACGATACGCTCCCAATATACTACTACATATCA 60
QY 246 ACCTTTTGCAGTGCAGAAAAGTACGCTGT 274
|||
Db 61 ACTTTTTCATTACAAAAAGTTTCATTT 89

RESULT 10

US-09-896-888A-15
; Sequence 15, Application US/09896888A
; Patent No. US20020116723A1
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/896,888A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/048,911
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/049,946

; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 88
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fragment of
; OTHER INFORMATION: promoter sequence of the AcMNPV ien gene
US-09-896-888A-15

Query Match 7.6%; Score 42.6; DB 9; Length 88;
Best Local Similarity 74.0%; Pred. No. 0.0016;
Matches 54; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 112 AACTTTTTTGCACTGCAAAAAAACACGCTTTTGACGCGGCCCATACATAGTACAAACT 171
|||
Db 16 AATTTTTTGCAATGCAAAAAAAGTTCACTTTTGCTGACACTCCCATATACAGTACAATCT 75
QY 172 CTACGTTTCGTAG 184
|||
Db 76 CTACAAATCGTAG 88

RESULT 11

US-10-473-126-199/c
; Sequence 199, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473,126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 199
; LENGTH: 6289
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-199

Query Match 7.0%; Score 39.6; DB 18; Length 6289;
Best Local Similarity 53.2%; Pred. No. 0.13;
Matches 84; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 112 AACTTTTTTGCACTGCAAAAAAACACGCTTTTGACGCGGCCCATACATAGTACAAACT 171
|||
Db 4896 AATTTTATTAACCAATATAAATCTATTATAAATCCACTACCAATATCATTTCAAAA 4837
QY 172 CTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCCAATACA 231
|||
Db 4836 TAATAATTTTAAACATTTATTATAAAAAATCGTCAATTTCTTTTACTCAAAAAATA 4777
QY 232 CTACACACATTAACCTTTTTCGAGTGCAGAAAAGTA 269
|||
Db 4776 ATCAACTTTCTCTCTCTATATAAAAAACGAAAAAATA 4739

RESULT 12

US-10-221-714A-223/c
; Sequence 223, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIERENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005

```

; CURRENT APPLICATION NUMBER: US/10/221-714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 223
; LENGTH: 9289
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-223

Query Match          7.0%; Score 39.6; DB 17; Length 9289;
Best Local Similarity 53.2%; Pred. No. 0.16;
Matches 84; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 112 AACTTTTTCGACTGCAAAAACACGCTTTTTCGACGGGGCCCATACATAGTACAAACT 171
Db 7896 AATTTATTATATACCAATAAATACTATTATAATCCACTACCAAAATATCATTTCAAAA 7837

QY 172 CTACGTTTCGTAGACTATTATTTATACATAAATAGTCTACACCGTTGTATACGCTCCAAATACA 231
Db 7836 TAATAATTTTAAACATTTATTATATAAATAATCGTCACAAATCTTTTACTCAAAAAATA 7777

QY 232 CTACACACATTTGAACCTTTTTCGACGTGCAAAAAAGTA 269
Db 7776 ATACAACCTTCTCTCTCCCTATATAAAGCAAAAAATA 7739

RESULT 13
US-10-473-126-345/c
; Sequence 345, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473,126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 345
; LENGTH: 6289
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-345

Query Match          6.7%; Score 38; DB 18; Length 6289;
Best Local Similarity 52.5%; Pred. No. 0.45;
Matches 83; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 112 AACTTTTTCGACTGCAAAAACACGCTTTTTCGACGGGGCCCATACATAGTACAAACT 171
Db 4896 AATTTATTATATACCAATAAATACTATTATAATCCACTACCAAAATATCATTTCAAAA 4837

QY 172 CTACGTTTCGTAGACTATTATTTATACATAAATAGTCTACACCGTTGTATACGCTCCAAATACA 231
Db 4836 TAATAATTTTAAACATTTATTATATAAATAATCATCACAATCTTTTACTCAAAAAATA 4777

QY 232 CTACCACATTTGAACCTTTTTCGACGTGCAAAAAAGTA 269

; CURRENT APPLICATION NUMBER: US/10/221-714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 223
; LENGTH: 9289
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-223

Query Match          7.0%; Score 39.6; DB 17; Length 9289;
Best Local Similarity 53.2%; Pred. No. 0.16;
Matches 84; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 112 AACTTTTTCGACTGCAAAAACACGCTTTTTCGACGGGGCCCATACATAGTACAAACT 171
Db 7896 AATTTATTATATACCAATAAATACTATTATAATCCACTACCAAAATATCATTTCAAAA 7837

QY 172 CTACGTTTCGTAGACTATTATTTATACATAAATAGTCTACACCGTTGTATACGCTCCAAATACA 231
Db 7836 TAATAATTTTAAACATTTATTATATAAATAATCGTCACAAATCTTTTACTCAAAAAATA 7777

QY 232 CTACACACATTTGAACCTTTTTCGACGTGCAAAAAAGTA 269
Db 7776 ATACAACCTTCTCTCTCCCTATATAAAGCAAAAAATA 7739

RESULT 14
US-10-398-221-1528/c
; Sequence 1528, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1528
; LENGTH: 986
; TYPE: DNA
; ORGANISM: Listeria monocytogenes-4B
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide a,g,c or t/u
US-10-398-221-1528

Query Match          6.3%; Score 35.4; DB 17; Length 986;
Best Local Similarity 56.4%; Pred. No. 1.3;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 155 CATACATAGTACAACTCTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTG 214
Db 318 CAGATTTTAGAAAACATTTTATCGTTCGCCAGGAAAATTTAAATATCAAGAGCACTCTGTTG 259

QY 215 TATACGCTCCAAATACACTACACACATTTGACCTTTTTCGAGTGCMAAAAAAGTAGC 271
Db 258 CAGAACATTCATATAAAGTTACATTCATTTTCGTTGCTGTAGAGGAG 202

RESULT 15
US-10-398-221-3152/c
; Sequence 3152, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3152
; LENGTH: 1549
; TYPE: DNA
; ORGANISM: Listeria monocytogenes 4b
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u
US-10-398-221-3152

Query Match          6.3%; Score 35.4; DB 17; Length 1549;
Best Local Similarity 56.4%; Pred. No. 1.6;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 155 CATACATAGTACAACTCTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTG 214
Db 318 CAGATTTTAGAAAACATTTTATCGTTCGCCAGGAAAATTTAAATATCAAGAGCACTCTGTTG 259

QY 215 TATACGCTCCAAATACACTACACACATTTGACCTTTTTCGAGTGCMAAAAAAGTAGC 271
Db 258 CAGAACATTCATATAAAGTTACATTCATTTTCGTTGCTGTAGAGGAG 202

RESULT 16
US-10-398-221-1528/c
; Sequence 1528, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1528
; LENGTH: 986
; TYPE: DNA
; ORGANISM: Listeria monocytogenes-4B
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide a,g,c or t/u
US-10-398-221-1528

Query Match          6.3%; Score 35.4; DB 17; Length 986;
Best Local Similarity 56.4%; Pred. No. 1.3;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 155 CATACATAGTACAACTCTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTG 214
Db 318 CAGATTTTAGAAAACATTTTATCGTTCGCCAGGAAAATTTAAATATCAAGAGCACTCTGTTG 259

QY 215 TATACGCTCCAAATACACTACACACATTTGACCTTTTTCGAGTGCMAAAAAAGTAGC 271
Db 258 CAGAACATTCATATAAAGTTACATTCATTTTCGTTGCTGTAGAGGAG 202
```

Matches	66;	Conservative	0;	Mismatches	51;	Indels	0;	Gaps	0;
Qy	155	CATACATAGTACAAACTCTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTG	214						
Db	881	CAGATTGAGAAAACATTTATCGTTGCCCGAGAAATTTAAATATCAAGAGCACTCTGTG	822						
Qy	215	TATACGCTCCAAATACACTACACACATTTGAACCTTTTTCAGTGCACAAAAAGTACG	271						
Db	821	CAGAACATTCATATAAAGTTACATCCATTGCTCAATTTTTCGGTGTCTAGAGGAG	765						

Search completed: May 10, 2005, 10:28:53
 Job time : 1600.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 04:33:07 ; Search time 4668.86 Seconds
(without alignments)
4931.434 Million cell updates/sec

Title: US-09-896-888A-1

Perfect score: 564

Sequence: 1 catgatgataacaatgat.....tggtacaggacacaacatg 564

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA Main:*

1: /cgn2_6/ptodata/1/pna/PCUTS1_COMB.seq:*

2: /cgn2_6/ptodata/1/pna/PCUTS2_COMB.seq:*

3: /cgn2_6/ptodata/1/pna/PCUTS_COMB.seq:*

4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*

5: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*

6: /cgn2_6/ptodata/1/pna/US08_COMB.seq:*

7: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*

8: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*

9: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*

10: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*

11: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*

12: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*

13: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*

14: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*

15: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*

16: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*

17: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*

18: /cgn2_6/ptodata/1/pna/US092A_COMB.seq:*

19: /cgn2_6/ptodata/1/pna/US092B_COMB.seq:*

20: /cgn2_6/ptodata/1/pna/US093A_COMB.seq:*

21: /cgn2_6/ptodata/1/pna/US093B_COMB.seq:*

22: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*

23: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*

24: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*

25: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*

26: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*

27: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*

28: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*

29: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*

30: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*

31: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*

32: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*

33: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*

34: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*

35: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*

36: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*

37: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*

38: /cgn2_6/ptodata/1/pna/US098D_COMB.seq:*

39: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*

40: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*

41: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*

42: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*

43: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*

44: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:*

45: /cgn2_6/ptodata/1/pna/US099G_COMB.seq:*

46: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*

47: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*

48: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*

49: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*

50: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*

51: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*

52: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*

53: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*

54: /cgn2_6/ptodata/1/pna/US104A_COMB.seq:*

55: /cgn2_6/ptodata/1/pna/US104B_COMB.seq:*

56: /cgn2_6/ptodata/1/pna/US105A_COMB.seq:*

57: /cgn2_6/ptodata/1/pna/US105B_COMB.seq:*

58: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*

59: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*

60: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*

61: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*

62: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*

63: /cgn2_6/ptodata/1/pna/US108A_COMB.seq:*

64: /cgn2_6/ptodata/1/pna/US108B_COMB.seq:*

65: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*

66: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*

67: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*

68: /cgn2_6/ptodata/1/pna/US110_COMB.seq:*

69: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*

70: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*

71: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*

72: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*

73: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*

74: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*

75: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*

76: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*

77: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*

78: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*

79: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*

80: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*

81: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*

82: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*

83: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*

84: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*

85: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*

86: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*

87: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*

88: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*

89: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*

90: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*

91: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*

92: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq:*

93: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq:*

94: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*

95: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*

96: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*

97: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*

98: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*

99: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*

100: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*

101: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*

102: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*

103: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*

104: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*

105: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*

106: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*

107: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*

108: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*

109: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*

110: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*

111: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*

112: /cgn2_6/ptodata/1/pna/US6042_COMB.seq:*

113: /cgn2_6/ptodata/1/pna/US6043_COMB.seq:*

114: /cgn2_6/ptodata/1/pna/US6044_COMB.seq:*

115: /cgn2_6/ptodata/1/pna/US6045_COMB.seq:*

116: /cgn2_6/ptodata/1/pna/US6046_COMB.seq:*

117: /cgn2_6/ptodata/1/pna/US6047_COMB.seq:
 118: /cgn2_6/ptodata/1/pna/US6048_COMB.seq:
 119: /cgn2_6/ptodata/1/pna/US6049_COMB.seq:
 120: /cgn2_6/ptodata/1/pna/US6050_COMB.seq:
 121: /cgn2_6/ptodata/1/pna/US6051_COMB.seq:
 122: /cgn2_6/ptodata/1/pna/US6052_COMB.seq:
 123: /cgn2_6/ptodata/1/pna/US6053_COMB.seq:
 124: /cgn2_6/ptodata/1/pna/US6054_COMB.seq:
 125: /cgn2_6/ptodata/1/pna/US6055_COMB.seq:
 126: /cgn2_6/ptodata/1/pna/US6056_COMB.seq:
 127: /cgn2_6/ptodata/1/pna/US6057_COMB.seq:
 128: /cgn2_6/ptodata/1/pna/US6058_COMB.seq:
 129: /cgn2_6/ptodata/1/pna/US6059_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	564	100.0	564	16	US-09-048-911-1	Sequence 1, Appli
2	564	100.0	564	38	US-09-896-888-1	Sequence 1, Appli
3	564	100.0	564	38	US-09-896-888A-1	Sequence 1, Appli
4	548	97.2	560	2	PCT-US03-22437-126	Sequence 126, App
5	548	97.2	560	58	US-10-622-088-126	Sequence 126, App
6	548	97.2	2773	51	US-10-295-074-60	Sequence 60, Appl
7	548	97.2	2773	63	US-10-846-911-60	Sequence 60, Appl
8	546.4	96.9	5038	2	PCT-US03-22437-89	Sequence 89, Appl
9	546.4	96.9	5038	58	US-10-622-088-89	Sequence 89, Appl
10	462	81.9	462	16	US-09-048-911-14	Sequence 14, Appl
11	462	81.9	462	38	US-09-896-888-14	Sequence 14, Appl
12	462	81.9	462	38	US-09-896-888A-14	Sequence 14, Appl
13	62	11.0	147	2	PCT-US03-22437-127	Sequence 127, App
14	62	11.0	147	58	US-10-622-088-127	Sequence 127, App
15	62	11.0	325	58	US-10-622-088-149	Sequence 149, App
16	57	10.1	92	16	US-09-048-911-16	Sequence 16, Appl
17	57	10.1	92	38	US-09-896-888-16	Sequence 16, Appl
18	57	10.1	92	38	US-09-896-888A-16	Sequence 16, Appl
19	42.6	7.6	88	16	US-09-048-911-15	Sequence 15, Appl
20	42.6	7.6	88	38	US-09-896-888-15	Sequence 15, Appl
21	42.6	7.6	88	38	US-09-896-888A-15	Sequence 15, Appl
22	39.6	7.0	6289	55	US-10-473-126-199	Sequence 199, App
23	39.6	7.0	9289	50	US-10-221-714B-223	Sequence 223, App
24	39.6	7.0	9289	50	US-10-221-714B-223	Sequence 223, App
25	38	6.7	6289	55	US-10-473-126-345	Sequence 345, App
26	36.4	6.5	505	58	US-10-664-025-15639	Sequence 15639, A
27	36.4	6.5	505	83	US-60-147-499-15639	Sequence 15639, A
28	36.2	6.4	425	23	US-09-521-640-252142	Sequence 252142, A
29	36.2	6.4	425	83	US-60-140-769-40512	Sequence 40512, A
30	36.2	6.4	2000	38	US-09-887-272A-5263	Sequence 5263, Ap
31	36	6.4	485	23	US-09-521-640-223561	Sequence 223561, A
32	35.8	6.3	187832	23	US-09-528-237A-1796	Sequence 1796, Ap
33	35.6	6.3	491	26	US-09-572-409-36238	Sequence 36238, A
34	35.6	6.3	514	36	US-09-834-366-36089	Sequence 36089, A
35	35.6	6.3	514	88	US-60-197-873-36089	Sequence 36089, A
36	35.6	6.3	877	50	US-10-238-659-475	Sequence 475, App
37	35.4	6.3	986	53	US-10-398-221-1528	Sequence 1528, Ap
38	35.4	6.3	1549	53	US-10-398-221-3152	Sequence 3152, Ap
39	35.2	6.2	367	46	US-10-020-338-4425	Sequence 4425, Ap
40	35	6.2	1254	51	US-10-266-090-35243	Sequence 35243, A
41	35	6.2	1852	69	US-60-606-062-4748	Sequence 4748, Ap
42	34.4	6.1	375	23	US-09-521-640-32406	Sequence 32406, A
43	34.4	6.1	375	26	US-09-552-087B-11171	Sequence 11171, A
44	34.4	6.1	375	26	US-09-552-087B-11171	Sequence 11171, A
45	34.4	6.1	855	28	US-09-621-011-89	Sequence 89, Appl

ALIGNMENTS

RESULT 1
 US-09-048-911-1
 ; Sequence 1, Application US/09048911
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of British Columbia
 ; TITLE OF INVENTION: Insect Expression Vectors
 ; FILE REFERENCE: 80021-44
 ; CURRENT APPLICATION NUMBER: US/09/048,911
 ; CURRENT FILING DATE: 1998-03-26
 ; EARLIER APPLICATION NUMBER: US 60/049,946
 ; EARLIER FILING DATE: 1997-03-27
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 564
 ; TYPE: DNA
 ; ORGANISM: Orgyia pseudotsugata
 US-09-048-911-1

Query Match 100.0%; Score 564; DB 16; Length 564;
 Best Local Similarity 100.0%; Pred. No. 5.7e-171;
 Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CATGATGATAAACAATGCTATGGTGTCTTAATCTGCTTCAACAACAATCTCTGTAACGTG 60
 Db 1 CATGATGATAAACAATGCTATGGTGTCTTAATCTGCTTCAACAACAATCTCTGTAACGTG 60
 QY 61 TTTTCATGTTTGGCAACAAGCAGCTTTTATCTCGTGGCTCCCAACCACTTTT 120
 Db 61 TTTTCATGTTTGGCAACAAGCAGCTTTTATCTCGTGGCTCCCAACCACTTTT 120
 QY 121 GCATGCAAAAAACAGCGCTTTTGACGCGGGGCGCATACATAGTACAACTCTACGTTTC 180
 Db 121 GCATGCAAAAAACAGCGCTTTTGACGCGGGGCGCATACATAGTACAACTCTACGTTTC 180
 QY 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATAGCTCCAAATACACTACCACAC 240
 Db 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATAGCTCCAAATACACTACCACAC 240
 QY 241 ATTGAACCTTTTTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 300
 Db 241 ATTGAACCTTTTTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 300
 QY 301 GGTGCGGCTCTGTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 360
 Db 301 GGTGCGGCTCTGTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 360
 QY 361 GACAGGACGCCAGCTTCTCTGT 420
 Db 361 GACAGGACGCCAGCTTCTCTGT 420
 QY 421 GGACGCGCTTCATATACGCGCGCTTATCTCATCGCGGTGACCGGACACGAGGCGCC 480
 Db 421 GGACGCGCTTCATATACGCGCGCTTATCTCATCGCGGTGACCGGACACGAGGCGCC 480
 QY 481 GTCCCGCTTATCGCGCTTAAATACAGCCCGCAACGATCTGTTAAACACAGTTGAACAG 540
 Db 481 GTCCCGCTTATCGCGCTTAAATACAGCCCGCAACGATCTGTTAAACACAGTTGAACAG 540
 QY 541 CATCTGTTACAGGACACCAATG 564
 Db 541 CATCTGTTACAGGACACCAATG 564

RESULT 2
 US-09-896-888-1
 ; Sequence 1, Application US/09896888
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of British Columbia
 ; TITLE OF INVENTION: Insect Expression Vectors
 ; FILE REFERENCE: 80021-44
 ; CURRENT APPLICATION NUMBER: US/09/896,888
 ; CURRENT FILING DATE: 2001-06-29

```
; PRIOR APPLICATION NUMBER: US 09/048,911
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: US 60/049,946
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Orgyia pseudotsugata
US-09-896-888-1

Query Match      100.0%; Score 564; DB 38; Length 564;
Best Local Similarity 100.0%; Pred. No. 5.7e-171;
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGATGATAAACAAGTATGGTCTAAATGCTTCAACAACAATTCGTGCAACTGTG 60
   |||
Db 1 CATGATGATAAACAAGTATGGTCTAAATGCTTCAACAACAATTCGTGCAACTGTG 60
   |||
QY 61 TTTTCATGTTTGGCAACAAGCACCCTTTATCTCGGTGGCCTCCCAACCAACTTTTTT 120
   |||
Db 61 TTTTCATGTTTGGCAACAAGCACCCTTTATCTCGGTGGCCTCCCAACCAACTTTTTT 120
   |||
QY 121 GCACTGCAAAAAAACAAGCTTTTGCAGCGGGGCCATACATAGTACAACTCTACGTTTC 180
   |||
Db 121 GCACTGCAAAAAAACAAGCTTTTGCAGCGGGGCCATACATAGTACAACTCTACGTTTC 180
   |||
QY 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCACAC 240
   |||
Db 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCACAC 240
   |||
QY 241 ATTGAACCTTTTTCAGAGTGCAAAAAAGTAGTGTGCGGAGTACGTAGCGCGGCTTATC 300
   |||
Db 241 ATTGAACCTTTTTCAGAGTGCAAAAAAGTAGTGTGCGGAGTACGTAGCGCGGCTTATC 300
   |||
QY 301 GGGTCGCGTCTGCTCAGTACGAATCACAATTATCGGACCGGACGAGTGTCTTATCGT 360
   |||
Db 301 GGGTCGCGTCTGCTCAGTACGAATCACAATTATCGGACCGGACGAGTGTCTTATCGT 360
   |||
QY 361 GACAGAGCGCAGCTTCTGTTGCTTAACCGCAGCGGACGCAACTCTTATCGGAACA 420
   |||
Db 361 GACAGAGCGCAGCTTCTGTTGCTTAACCGCAGCGGACGCAACTCTTATCGGAACA 420
   |||
QY 421 GGAAGCGCTTCCATATCAGCGCGGCTTATCTCATGCGGTGACCGGACGAGGCGGCC 480
   |||
Db 421 GGAAGCGCTTCCATATCAGCGCGGCTTATCTCATGCGGTGACCGGACGAGGCGGCC 480
   |||
QY 481 GTCCCGCTTATCGCGCTATAAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAG 540
   |||
Db 481 GTCCCGCTTATCGCGCTATAAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAG 540
   |||
QY 541 CATCTGTTACAGGACACACATG 564
   |||
Db 541 CATCTGTTACAGGACACACATG 564
   |||
```

```
RESULT 3
US-09-896-888a-1
; Sequence 1, Application US/09896888A
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/896,888A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/048,911
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/049,946
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
```

```
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Orgyia pseudotsugata
US-09-896-888A-1

Query Match      100.0%; Score 564; DB 38; Length 564;
Best Local Similarity 100.0%; Pred. No. 5.7e-171;
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGATGATAAACAAGTATGGTCTAAATGCTTCAACAACAATTCGTGCAACTGTG 60
   |||
Db 1 CATGATGATAAACAAGTATGGTCTAAATGCTTCAACAACAATTCGTGCAACTGTG 60
   |||
QY 61 TTTTCATGTTTGGCAACAAGCACCCTTTATCTCGGTGGCCTCCCAACCAACTTTTTT 120
   |||
Db 61 TTTTCATGTTTGGCAACAAGCACCCTTTATCTCGGTGGCCTCCCAACCAACTTTTTT 120
   |||
QY 121 GCACTGCAAAAAAACAAGCTTTTGCAGCGGGGCCATACATAGTACAACTCTACGTTTC 180
   |||
Db 121 GCACTGCAAAAAAACAAGCTTTTGCAGCGGGGCCATACATAGTACAACTCTACGTTTC 180
   |||
QY 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCACAC 240
   |||
Db 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCACAC 240
   |||
QY 241 ATTGAACCTTTTTCAGAGTGCAAAAAAGTAGTGTGCGGAGTACGTAGCGCGGCTTATC 300
   |||
Db 241 ATTGAACCTTTTTCAGAGTGCAAAAAAGTAGTGTGCGGAGTACGTAGCGCGGCTTATC 300
   |||
QY 301 GGGTCGCGTCTGCTCAGTACGAATCACAATTATCGGACCGGACGAGTGTCTTATCGT 360
   |||
Db 301 GGGTCGCGTCTGCTCAGTACGAATCACAATTATCGGACCGGACGAGTGTCTTATCGT 360
   |||
QY 361 GACAGAGCGCAGCTTCTGTTGCTTAACCGCAGCGGACGCAACTCTTATCGGAACA 420
   |||
Db 361 GACAGAGCGCAGCTTCTGTTGCTTAACCGCAGCGGACGCAACTCTTATCGGAACA 420
   |||
QY 421 GGAAGCGCTTCCATATCAGCGCGGCTTATCTCATGCGGTGACCGGACGAGGCGGCC 480
   |||
Db 421 GGAAGCGCTTCCATATCAGCGCGGCTTATCTCATGCGGTGACCGGACGAGGCGGCC 480
   |||
QY 481 GTCCCGCTTATCGCGCTATAAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAG 540
   |||
Db 481 GTCCCGCTTATCGCGCTATAAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAG 540
   |||
QY 541 CATCTGTTACAGGACACACATG 564
   |||
Db 541 CATCTGTTACAGGACACACATG 564
   |||
```

```
RESULT 4
PCT-US03-22437-126
; Sequence 126, Application PC/TUS0322437
; GENERAL INFORMATION:
; APPLICANT: Invitrogen Corporation
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.545PC07
; CURRENT APPLICATION NUMBER: PCT/US03/22437
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 126
; LENGTH: 560
```

```
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: OpiB2 promoter sequence
PCT-US03-22437-126

Query Match          97.2%; Score 548; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 8.6e-166;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGATGATAAACAATGATGCTTAATGCTTCAACAACAATTCGTGAACTGTG 60
Db 5 CATGATGATAAACAATGATGCTTAATGCTTCAACAACAATTCGTGAACTGTG 64

QY 61 TTTTCATGTTGGCAACAGCAGCTTTTATCTCGTGGCTCCCAACAACAATTCGTG 120
Db 65 TTTTCATGTTGGCAACAGCAGCTTTTATCTCGTGGCTCCCAACAACAATTCGTG 124

QY 121 GCACTGCAAAAAAACAAGCCTTTTTCACGCGGCGCCATACATAGTACAAACTCTAGTTTC 180
Db 125 GCACTGCAAAAAAACAAGCCTTTTTCACGCGGCGCCATACATAGTACAAACTCTAGTTTC 184

QY 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCAAC 240
Db 185 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCAAC 244

QY 241 ATTGAACCTTTTTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 300
Db 245 ATTGAACCTTTTTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 304

QY 301 GGGTCCGCTTCTGTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 360
Db 305 GGGTCCGCTTCTGTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 364

QY 361 GACAGACCCAGCTTCTGTTGCTAAACCGCAGCGGACCAACTCTCTATCGGAACA 420
Db 365 GACAGACCCAGCTTCTGTTGCTAAACCGCAGCGGACCAACTCTCTATCGGAACA 424

QY 421 GGACGGCTCCATATCAGCGGGGTTATCTCATCGCGGTACCGGACAGAGCGGCC 480
Db 425 GGACGGCTCCATATCAGCGGGGTTATCTCATCGCGGTACCGGACAGAGCGGCC 484

QY 481 GTCCGCTTATCGCGCTATAAATACAGCCCGCAAGATCTGGTAAACACAGTTGAAACAG 540
Db 485 GTCCGCTTATCGCGCTATAAATACAGCCCGCAAGATCTGGTAAACACAGTTGAAACAG 544

QY 541 CATCTGTT 548
Db 545 CATCTGTT 552
```

```
RESULT 5
US-10-622-088-126
; Sequence 126, Application US/10622088
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
```

```
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 126
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: OpiB2 promoter sequence
US-10-622-088-126
```

```
Query Match          97.2%; Score 548; DB 58; Length 560;
Best Local Similarity 100.0%; Pred. No. 8.6e-166;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGATGATAAACAATGATGCTTAATGCTTCAACAACAATTCGTGAACTGTG 60
Db 5 CATGATGATAAACAATGATGCTTAATGCTTCAACAACAATTCGTGAACTGTG 64

QY 61 TTTTCATGTTGGCAACAGCAGCTTTTATCTCGTGGCTCCCAACAACAATTCGTG 120
Db 65 TTTTCATGTTGGCAACAGCAGCTTTTATCTCGTGGCTCCCAACAACAATTCGTG 124

QY 121 GCACTGCAAAAAAACAAGCCTTTTTCACGCGGCGCCATACATAGTACAAACTCTAGTTTC 180
Db 125 GCACTGCAAAAAAACAAGCCTTTTTCACGCGGCGCCATACATAGTACAAACTCTAGTTTC 184

QY 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCAAC 240
Db 185 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCAAC 244

QY 241 ATTGAACCTTTTTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 300
Db 245 ATTGAACCTTTTTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 304

QY 301 GGGTCCGCTTCTGTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 360
Db 305 GGGTCCGCTTCTGTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 364

QY 361 GACAGACCCAGCTTCTGTTGCTAAACCGCAGCGGACCAACTCTCTATCGGAACA 420
Db 365 GACAGACCCAGCTTCTGTTGCTAAACCGCAGCGGACCAACTCTCTATCGGAACA 424

QY 421 GGACGGCTCCATATCAGCGGGGTTATCTCATCGCGGTACCGGACAGAGCGGCC 480
Db 425 GGACGGCTCCATATCAGCGGGGTTATCTCATCGCGGTACCGGACAGAGCGGCC 484

QY 481 GTCCGCTTATCGCGCTATAAATACAGCCCGCAAGATCTGGTAAACACAGTTGAAACAG 540
Db 485 GTCCGCTTATCGCGCTATAAATACAGCCCGCAAGATCTGGTAAACACAGTTGAAACAG 544

QY 541 CATCTGTT 548
Db 545 CATCTGTT 552
```

```
RESULT 6
US-10-295-074-60
; Sequence 60, Application US/10295074
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: P1013DK00
; CURRENT APPLICATION NUMBER: US/10/295,074
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 2773
; TYPE: DNA
```

```

; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p2Op2F expression vector for insect cells
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (561)..(566)
; OTHER INFORMATION: HindIII site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (573)..(578)
; OTHER INFORMATION: Aval site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (586)..(591)
; OTHER INFORMATION: EcoRI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (593)..(598)
; OTHER INFORMATION: BamHI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (625)..(630)
; OTHER INFORMATION: ClaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (629)..(634)
; OTHER INFORMATION: ClaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (1156)..(1161)
; OTHER INFORMATION: ApaLI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2128)..(2133)
; OTHER INFORMATION: PstI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2204)..(2209)
; OTHER INFORMATION: NcoI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2284)..(2289)
; OTHER INFORMATION: Aval site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2294)..(2299)
; OTHER INFORMATION: Aval, SmaI, and XmaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2551)..(2556)
; OTHER INFORMATION: ApaLI site
;
US-10-295-074-60

```

```

Query Match      97.2%; Score 548; DB 51; Length 2773;
Best Local Similarity 100.0%; Pred. No. 1.7e-165;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CATGATGATAAACAATGTATGGTGTAAATGCTTCAACAAACAATTCGTGAACTGTG 60
DB      5  CATGATGATAAACAATGTATGGTGTAAATGCTTCAACAAACAATTCGTGAACTGTG 64
QY     61  TTTTCATGTTTGGCAACAACACCTTTATATCTCGGTGGCTCCACCACCAACTTTTTT 120
DB     65  TTTTCATGTTTGGCAACAACACCTTTATATCTCGGTGGCTCCACCACCAACTTTTTT 124
QY    121  GCACGTCAAAAAACACGCTTTTGCACGGGGCCCATACATAGTACAAACTCTACGTTTC 180
DB    125  GCACGTCAAAAAACACGCTTTTGCACGGGGCCCATACATAGTACAAACTCTACGTTTC 184
QY    181  GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCAC 240
DB    185  GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCAC 244

```

```

QY      241  ATTGAACCTTTTTCAGTGCAAAAAAGTAGCTGCGGAGTACGCTAGGCGCGCTTATC 300
DB     245  ATTGAACCTTTTTCAGTGCAAAAAAGTAGCTGCGGAGTACGCTAGGCGCGCTTATC 304
QY     301  GGGTCGCGCTCTGTACAGTACGAATCACATTATCGGACCGGACGAGTGTGTTATCGT 360
DB     305  GGGTCGCGCTCTGTACAGTACGAATCACATTATCGGACCGGACGAGTGTGTTATCGT 364
QY     361  GACAGGACGCGACGCTTCTGTGTGTAAACCGGACGCGGACGCAACTCCTTATCGGAACA 420
DB     365  GACAGGACGCGACGCTTCTGTGTGTAAACCGGACGCGGACGCAACTCCTTATCGGAACA 424
QY     421  GAGCGCGCTTCATATCAGCGCGCGCTTATCTCATGGCGTGACCGGACGAGGCGCCC 480
DB     425  GAGCGCGCTTCATATCAGCGCGCGCTTATCTCATGGCGTGACCGGACGAGGCGCCC 484
QY     481  GTCCCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCTGTTAAACACAGTTGAACAG 540
DB     485  GTCCCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCTGTTAAACACAGTTGAACAG 544
QY     541  CATCTGTT 548
DB     545  CATCTGTT 552

RESULT 7
US-10-846-911-60
; Sequence 60, Application US/10846911
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen
; APPLICANT: NIELSEN, Finn Stausholm
; APPLICANT: BRATT, Tomas
; APPLICANT: VOLDOR, Bjorn
; APPLICANT: MOURITSEN, Soren
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: 674542-2018
; CURRENT APPLICATION NUMBER: US/10/846,911
; CURRENT FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: PCT/DK02/00764
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: 60/331,575
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: PA 2001 01702
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 2773
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p2Op2F expression vector for insect cells
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (561)..(566)
; OTHER INFORMATION: HindIII site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (573)..(578)
; OTHER INFORMATION: Aval site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (586)..(591)
; OTHER INFORMATION: EcoRI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (593)..(598)
; OTHER INFORMATION: BamHI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (625)..(630)
; OTHER INFORMATION: ClaI site
; FEATURE:

```

```

; NAME/KEY: misc_recomb
; LOCATION: (629)..(634)
; OTHER INFORMATION: Clai site
; FEATURE:
;
; NAME/KEY: misc_recomb
; LOCATION: (1156)..(1161)
; OTHER INFORMATION: Apali site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2128)..(2133)
; OTHER INFORMATION: Psti site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2204)..(2209)
; OTHER INFORMATION: NcoI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2284)..(2289)
; OTHER INFORMATION: Aval site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2294)..(2299)
; OTHER INFORMATION: Aval, SmaI
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2551)..(2556)
; OTHER INFORMATION: Apali site
; US-10-846-911-60

```

Query Match	97.2%	Score 548	DB 63	Length 2773
Best Local Similarity	100.0%	Pred. No. 1.7e-165		
Matches 548	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	CATGATGATAAACAAATGTATGTCGTCTAATGTCCTTCAACAACAATTCGTGTGAACCTGTG	60	
Db	5	CATGATGATAAACAAATGTATGTCGTCTAATGTCCTTCAACAACAATTCGTGTGAACCTGTG	64	
QY	61	TTTTTCATGTTTGGCCAAACAGCAGCTTTATACTCGGTGGGCTCCCAACACCAACTTTTTT	120	
Db	65	TTTTTCATGTTTGGCCAAACAGCAGCTTTATACTCGGTGGGCTCCCAACACCAACTTTTTT	124	
QY	121	GCATGCAAAAAAAAACAGCTTTTGCACGGCGGCCCATACATAGTACAAACTCTACGTTTC	180	
Db	125	GCATGCAAAAAAAAACAGCTTTTGCACGGCGGCCCATACATAGTACAAACTCTACGTTTC	184	
QY	181	GTAGACTATTTTACATAAAATAGTCTACACCGTTGTATACGCTCCAAATACACTCCACAC	240	
Db	185	GTAGACTATTTTACATAAAATAGTCTACACCGTTGTATACGCTCCAAATACACTCCACAC	244	
QY	241	ATTGAACCTTTTTCGAGTCGAAAAAGTACGTCGCGCAGTCACGTAGCCGGCCTTATC	300	
Db	245	ATTGAACCTTTTTCGAGTCGAAAAAGTACGTCGCGCAGTCACGTAGCCGGCCTTATC	304	
QY	301	GGGTGCGGCTCTGTACGTCAGAAATCACATTTATCGGACCGGACGAGTGTGTCTATCGT	360	
Db	305	GGGTGCGGCTCTGTACGTCAGAAATCACATTTATCGGACCGGACGAGTGTGTCTATCGT	364	
QY	361	GACAGGACCGCAGCTTCTGTGTTGTAAACCGCAGCCGGACGGAACTCCTTATCGGAACA	420	
Db	365	GACAGGACCGCAGCTTCTGTGTTGTAAACCGCAGCCGGACGGAACTCCTTATCGGAACA	424	
QY	421	GGACGGCGCTCCATATCAGCCGCGCGTTATCTCATGCGGTGACCGGACACGAGGCGGCC	480	
Db	425	GGACGGCGCTCCATATCAGCCGCGCGTTATCTCATGCGGTGACCGGACACGAGGCGGCC	484	
QY	481	GTCGGCTTATCGGCTTAAATAACAGCCCGCAACGATCTGGTAAACACAGTTGAAACAG	540	
Db	485	GTCGGCTTATCGGCGCTTAAATAACAGCCCGCAACGATCTGGTAAACACAGTTGAAACAG	544	
QY	541	CATCTGTT 548		
Db	545	CATCTGTT 552		

```

RESULT 8
PCT-US03-22437-89
; Sequence 89, Application PC/TUS0322437
; GENERAL INFORMATION:
; APPLICANT: Invitrogen Corporation
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.545PC07
; CURRENT APPLICATION NUMBER: PCT/US03/22437
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 89
; LENGTH: 5038
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: piB/V5-His-DEST
PCT-US03-22437-89

```

Query Match	96.9%;	Score 546.4;	DB 2;	Length 5038;
Best Local Similarity	99.8%;	Pred. No. 7.3e-165;		
Matches 547;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	CATGATGATAACAAATGATCGTCTAATGTTGCTTCAACAACAATTCGTGTGAACGTGG	60	
Db	1			
QY	61	TTTTTCATGTTTGGCCAACAAGCACCTTTATCTACTCGTGGCCCTCCCAACCACTTTT	120	
Db	61	TTTTTCATGTTTGGCCAACAAGCACCTTTATCTACTCGTGGCCCTCCCAACCACTTTT	120	
QY	121	GCACGTGCAAAAAACACGCTTTTTCACGCGGGCCCATACATAGTACAAACTCTACGTTTC	180	
Db	121	GCACGTGCAAAAAACACGCTTTTTCACGCGGGCCCATACATAGTACAAACTCTACGTTTC	180	
QY	181	GTAGACTATTTTACATAAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCACAC	240	
Db	181	GTAGACTATTTTACATAAAATAGTCTACACCGTTGTATACGCTCCAAATACACTACCACAC	240	
QY	241	ATTGAACCTTTTTTGCGAGTGCAAAAAGTACGTGTGCGCAGTACGTAGCGCGCCCTTATC	300	
Db	241	ATTGAACCTTTTTTGCGAGTGCAAAAAGTACGTGTGCGCAGTACGTAGCGCGCCCTTATC	300	
QY	301	GGGTGCGGTGCTGTACGTACGAATCACATTATTCGGACCGGACGAGTGTGCTTATCGT	360	
Db	301	GGGTGCGGTGCTGTACGTACGAATCACATTATTCGGACCGGACGAGTGTGCTTATCGT	360	
QY	361	GACGAGCGCCAGCTTCTCTGTGTTGCTTAACCGCAGCGGACGCAACTCTTATCGGAACA	420	
Db	361	GACGAGCGCCAGCTTCTCTGTGTTGCTTAACCGCAGCGGACGCAACTCTTATCGGAACA	420	
QY	421	GGACGGGCTTCATATCAGCGCGCGTTATCTCATGCGGTGACCGGACACGAGGGCGCC	480	
Db	421	GGACGGGCTTCATATCAGCGCGCGTTATCTCATGCGGTGACCGGACACGAGGGCGCC	480	
QY	481	GTCCCGCTTATTCGGGCCCTATAAATACAGCCGCAACGATCTGTTAAACACAGTTTGAACAG	540	
Db	481	GTCCCGCTTATTCGGGCCCTATAAATACAGCCGCAACGATCTGTTAAACACAGTTTGAACAG	540	
QY	541	CATCTGTT	548	
Db	541	CATCTGTT	548	

RESULT 9

US-10-622-088-89
; Sequence 89, Application US/10622088
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622.088
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 89
; LENGTH: 5038
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pIB/V5-His-DEST
US-10-622-088-89

Query Match 96.9%; Score 546.4; DB 58; Length 5038;
Best Local Similarity 99.8%; Pred. No. 7.3e-165;
Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CATGATGATAAACAATGTATGGTCTAATGTGCTTCAACAAACAATTCCTGTAACGTG 60
DB 1 CATGATGATAAACAATGTATGGTCTAATGTGCTTCAACAAACAATTCCTGTAACGTG 60
QY 61 TTTTCATGTTTGGCAACAGCAGCTTTATCTAGTGGTGGCTCCACCACCACTTTT 120
DB 61 TTTTCATGTTTGGCAACAGCAGCTTTATCTAGTGGTGGCTCCACCACCACTTTT 120
QY 121 GCACTGCAAAAAACAGCCTTTTGACGGGGCCATACATAGTACAAACTCTACGTTTC 180
DB 121 GCACTGCAAAAAACAGCCTTTTGACGGGGCCATACATAGTACAAACTCTACGTTTC 180
QY 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACCGTCCAAATACACTACCAC 240
DB 181 GTAGACTATTTTACATAAATAGTCTACACCGTTGTATACCGTCCAAATACACTACCAC 240
QY 241 ATTGAACCTTTTGCAGTCGCAAAAAGTACGTGTCGAGTCAGTAGGGCGGCTTATC 300
DB 241 ATTGAACCTTTTGCAGTCGCAAAAAGTACGTGTCGAGTCAGTAGGGCGGCTTATC 300
QY 301 GGGTCGGCTCTGTGTCAGTACGAATACATTTATCGGACCGGACGAGTGTGTTATCGT 360
DB 301 GGGTCGGCTCTGTGTCAGTACGAATACATTTATCGGACCGGACGAGTGTGTTATCGT 360
QY 361 GACAGGACCGCAGCTTCTGTGTTGCTTAAACCGAGCGGACGCACTCTTATCGGAACA 420
DB 361 GACAGGACCGCAGCTTCTGTGTTGCTTAAACCGAGCGGACGCACTCTTATCGGAACA 420
QY 421 GGACGGCCTTCATATCAGCGGCGGCTTATCTATGCGGCTGACCGGACAGAGGGGCC 480
DB 421 GGACGGCCTTCATATCAGCGGCGGCTTATCTATGCGGCTGACCGGACAGAGGGGCC 480

QY 481 GTCCCGCTTATCGCGCTATAAATACAGCCCGCAACGATCTGTAAACACAGTTGAACAG 540
DB 481 GTCCCGCTTATCGCGCTATAAATACAGCCCGCAACGATCTGTAAACACAGTTGAACAG 540
QY 541 CATCTGTT 548
DB 541 CATCTGTT 548

RESULT 10

US-09-048-911-14
; Sequence 14, Application US/09048911
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/048,911
; CURRENT FILING DATE: 1998-03-26
; EARLIER APPLICATION NUMBER: US 60/049,946
; EARLIER FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Promoter
; OTHER INFORMATION: sequence of the OpMPV iε2 gene
US-09-048-911-14

Query Match 81.9%; Score 462; DB 16; Length 462;
Best Local Similarity 100.0%; Pred. No. 5.5e-138;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 102 CCCACACCAACTTTTTCGACTGCAAAAAACAGCTTTTCGACGCGGCGCCATACAT 161
DB 1 CCCACACCAACTTTTTCGACTGCAAAAAACAGCTTTTCGACGCGGCGCCATACAT 60
QY 162 AGTACAACTCTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGC 221
DB 61 AGTACAACTCTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGC 120
QY 222 TCCAAATACATACACACACATTGAACCTTTTTCAGTGCAAAAAAGTACGTGCGGAGT 281
DB 121 TCCAAATACATACACACATTGAACCTTTTTCAGTGCAAAAAAGTACGTGCGGAGT 180
QY 282 CACGTAGGCGGCGCTTATCGGTCGCTCTGTACGTACGAATCACTTATCGGACCGG 341
DB 181 CACGTAGGCGGCGCTTATCGGTCGCTCTGTACGTACGAATCACTTATCGGACCGG 240
QY 342 ACAGTGTGTTCTTATCTGTCAGGACGCGGCTTCTGTGTTGCTTAACCGGACCGGAC 401
DB 241 ACAGTGTGTTCTTATCTGTCAGGACGCGGCTTCTGTGTTGCTTAACCGGACCGGAC 300
QY 402 GCAACTCTTATCGGAACAGGACGCGCTCCATATCAGCGCGCGCTTATCTCTATCGCGT 461
DB 301 GCAACTCTTATCGGAACAGGACGCGCTCCATATCAGCGCGCGCTTATCTCTATCGCGT 360
QY 462 GACCGGACAGGCGCGCTTATCGCGCTTATCGCGCTTATCGCGCTTATCGCGCTTATCG 521
DB 361 GACCGGACAGGCGCGCTTATCGCGCTTATCGCGCTTATCGCGCTTATCGCGCTTATCG 420
QY 522 GGTAAACAGTGTGAACAGACTCTGTTACGCGGACCAACAT 563
DB 421 GGTAAACAGTGTGAACAGACTCTGTTACGCGGACCAACAT 462

RESULT 11

US-09-896-888-14
; Sequence 14, Application US/09896888
; GENERAL INFORMATION:

```
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/896,888
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/048,911
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: US 60/049,946
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 14
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Promoter
; OTHER INFORMATION: sequence of the OpNPV ie2 gene
US-09-896-888-14

Query Match      81.9%; Score 462; DB 38; Length 462;
Best Local Similarity 100.0%; Pred. No. 5.5e-138;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 CCCACACCAACTTTTTTGCACTGCAAAAAACACGCTTTTGCACGGGGCCCATACAT 161
Db 1 CCCACACCAACTTTTTTGCACTGCAAAAAACACGCTTTTGCACGGGGCCCATACAT 60

QY 162 AGTACAACTCTAGCTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGC 221
Db 61 AGTACAACTCTAGCTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGC 120

QY 222 TCCAAATACACTACACACATTTGCACTTTTGCAGTGCAGTCAAAAAAGTACGTGTCGGCAGT 281
Db 61 AGTACAACTCTAGCTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGC 120

QY 282 CACGTAGGCGGCGCTTATCGGCTCGCTCTGCTACAGTACGAATCACATTATCGGACCGG 341
Db 181 CACGTAGGCGGCGCTTATCGGCTCGCTCTGCTACAGTACGAATCACATTATCGGACCGG 240

QY 342 ACAGTGTGTCTTATCGTGACAGACGCCAGCTTCTGCTGTGCTAAACCGCAGCCGAC 401
Db 241 CACGTAGGCGGCGCTTATCGGCTCGCTCTGCTACAGTACGAATCACATTATCGGACCGG 240

QY 402 ACAGTGTGTCTTATCGTGACAGACGCCAGCTTCTGCTGTGCTAAACCGCAGCCGAC 461
Db 241 ACAGTGTGTCTTATCGTGACAGACGCCAGCTTCTGCTGTGCTAAACCGCAGCCGAC 300

QY 462 GCAACTCTTATCGGAACAGGACGCCCTTCCATATCAGCGCGCTTATCTCATCGCGT 461
Db 301 GCAACTCTTATCGGAACAGGACGCCCTTCCATATCAGCGCGCTTATCTCATCGCGT 360

QY 462 GACCGGACAGGCGCGCGCTTATCGCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCT 521
Db 361 GACCGGACAGGCGCGCGCTTATCGCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCT 420

QY 522 GGTAAACACAGTTGAAACAGCATCTGTTACAGCGACACAACAT 563
Db 421 GGTAAACACAGTTGAAACAGCATCTGTTACAGCGACACAACAT 462

RESULT 12
US-09-896-888A-14
; Sequence 14, Application US/09896888A
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/896,888A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/048,911
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/049,946
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 14
```

```
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Promoter
; OTHER INFORMATION: sequence of the OpNPV ie2 gene
US-09-896-888A-14

Query Match      81.9%; Score 462; DB 38; Length 462;
Best Local Similarity 100.0%; Pred. No. 5.5e-138;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 CCCACACCAACTTTTTTGCACTGCAAAAAACACGCTTTTGCACGGGGCCCATACAT 161
Db 1 CCCACACCAACTTTTTTGCACTGCAAAAAACACGCTTTTGCACGGGGCCCATACAT 60

QY 162 AGTACAACTCTAGCTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGC 221
Db 61 AGTACAACTCTAGCTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGC 120

QY 222 TCCAAATACACTACACACATTTGCACTTTTGCAGTGCAGTCAAAAAAGTACGTGTCGGCAGT 281
Db 121 TCCAAATACACTACACACATTTGCACTTTTGCAGTGCAGTCAAAAAAGTACGTGTCGGCAGT 180

QY 282 CACGTAGGCGGCGCTTATCGGCTCGCTCTGCTACAGTACGAATCACATTATCGGACCGG 341
Db 181 CACGTAGGCGGCGCTTATCGGCTCGCTCTGCTACAGTACGAATCACATTATCGGACCGG 240

QY 342 ACAGTGTGTCTTATCGTGACAGACGCCAGCTTCTGCTGTGCTAAACCGCAGCCGAC 401
Db 241 ACAGTGTGTCTTATCGTGACAGACGCCAGCTTCTGCTGTGCTAAACCGCAGCCGAC 300

QY 402 GCAACTCTTATCGGAACAGGACGCCCTTCCATATCAGCGCGCTTATCTCATCGCGT 461
Db 301 GCAACTCTTATCGGAACAGGACGCCCTTCCATATCAGCGCGCTTATCTCATCGCGT 360

QY 462 GACCGGACAGGCGCGCGCTTATCGCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCT 521
Db 361 GACCGGACAGGCGCGCGCTTATCGCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCT 420

QY 522 GGTAAACACAGTTGAAACAGCATCTGTTACAGCGACACAACAT 563
Db 421 GGTAAACACAGTTGAAACAGCATCTGTTACAGCGACACAACAT 462

RESULT 13
PCT-US03-22437-127
; Sequence 127, Application PC/TUS0322437
; GENERAL INFORMATION:
; APPLICANT: Invitrogen Corporation
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.545PC07
; CURRENT APPLICATION NUMBER: PCT/US03/22437
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 127
; LENGTH: 147
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombination region of pIB/V5-His-DEST
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (141)..(148)
; OTHER INFORMATION: n may be any nucleotide
PCT-US03-22437-127

Query Match      11.0%; Score 62; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 487 CTTATCGCGCTATAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAGCATCTG 546
Db 1 CTTATCGCGCTATAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAGCATCTG 60

QY 547 TT 548
Db 61 TT 62

RESULT 14
US-10-622-088-127
; Sequence 127, Application US/10622088
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 127
; LENGTH: 147
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombination region of pIB/V5-His-DEST
; NAME/KEY: misc feature
; LOCATION: (141)..(148)
; OTHER INFORMATION: n may be any nucleotide
US-10-622-088-127

Query Match      11.0%; Score 62; DB 58; Length 147;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 487 CTTATCGCGCTATAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAGCATCTG 546
Db 1 CTTATCGCGCTATAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAGCATCTG 60

QY 547 TT 548
Db 61 TT 62
```

RESULT 15
US-10-622-088-149

```
; Sequence 149, Application US/10622088
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth E.
; APPLICANT: Frank, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 149
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Recombination region of pIB/V5 His DEST
; NAME/KEY: misc feature
; LOCATION: (141)..(142)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (145)..(276)
US-10-622-088-149

Query Match      11.0%; Score 62; DB 58; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 487 CTTATCGCGCTATAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAGCATCTG 546
Db 1 CTTATCGCGCTATAATACAGCCCGCAACGATCTGGTAAACACAGTTGAACAGCATCTG 60

QY 547 TT 548
Db 61 TT 62

Search completed: May 10, 2005, 09:18:49
Job time : 4670.86 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 05:51:35 ; Search time 1390.39 Seconds
(without alignments)
834.449 Million cell updates/sec

Title: US-09-896-888A-1
Perfect score: 564
Sequence: 1 catgatgataacaatgtat.....tggtagcgacacacatg 564

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9398789 seqs, 102855566 residues

Total number of hits satisfying chosen parameters: 18797578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_NA_New.*

- 1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq2.*
- 3: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
- 8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq2.*
- 9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq3.*
- 10: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq.*
- 11: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq2.*
- 12: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq3.*
- 13: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	548	97.2	2773	7	US-10-939-107-60
2	34	6.0	18977	10	Sequence 26311, A
C 3	33.8	6.0	600	11	US-11-097-143-26311
C 4	33.8	6.0	600	11	US-11-060-756-2178
C 5	33.8	6.0	600	11	US-11-060-756-2179
C 6	33.8	6.0	600	11	US-11-060-756-6450
C 7	33.8	6.0	600	11	US-11-060-756-6451
C 8	33.8	6.0	3057	2	PCT-US05-07748-6
C 9	33.6	6.0	1391	13	US-60-655-875-26427
C 10	33.2	5.9	43053	9	US-10-472-963-1098
C 11	32.4	5.7	1221	11	US-11-079-463-1097
C 12	32.4	5.7	1221	11	US-11-079-045-3281
C 13	32.4	5.7	1268	10	US-11-097-143-25442
C 14	32.4	5.7	1342	10	US-11-097-143-38495
C 15	32.4	5.7	3448	10	US-11-097-143-38494
C 16	32.4	5.7	3600	10	US-11-097-143-25441
C 17	32.4	5.7	3891	10	US-11-097-143-33844
C 18	32.2	5.7	599	9	US-10-972-079-15168
C 19	32.2	5.7	600	9	US-10-972-079-15169
C 20	32.2	5.7	1915	13	US-60-655-875-2139

C 21	32.2	5.7	22319	10	US-11-097-143-37270
C 22	31.6	5.6	1550	13	US-60-655-875-63563
C 23	31.6	5.6	1767	13	US-60-655-875-7011
C 24	31.6	5.6	7833	8	US-10-517-441-350
C 25	31.6	5.6	7833	8	US-10-517-441-624
C 26	31.4	5.6	1015	8	US-10-450-763-18397
C 27	31.4	5.6	20670	13	US-60-659-397-12262
C 28	31	5.5	549	11	US-11-031-175-6433
C 29	31	5.5	1356	13	US-60-655-875-1528
C 30	31	5.5	12865	11	US-11-031-175-1048
C 31	31	5.5	45938	11	US-11-073-360-1642
C 32	31	5.5	71585	1	PCT-US04-31416-50
C 33	31	5.5	435302	13	US-60-660-322-2196
C 34	30.8	5.5	33249	10	US-11-097-143-20386
C 35	30.6	5.4	446	13	US-60-669-175-18851
C 36	30.6	5.4	637	9	US-10-703-032-68319
C 37	30.6	5.4	638	9	US-10-703-032-53281
C 38	30.6	5.4	3867	13	US-60-655-875-6428
C 39	30.6	5.4	9353	8	US-10-517-441-462
C 40	30.6	5.4	9353	8	US-10-517-441-736
C 41	30.6	5.4	10968	2	PCT-US05-07924-35
C 42	30.6	5.4	10968	11	US-11-075-185-35
C 43	30.6	5.4	21231	10	US-11-097-143-25744
C 44	30.6	5.4	25048	11	US-11-031-175-1239
C 45	30.6	5.4	78869	2	PCT-US05-07924-1

ALIGNMENTS

RESULT 1

US-10-939-107-60
; Sequence 60, Application US/10939107
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Hans Rudolf
; APPLICANT: Ebert, Bjørke
; APPLICANT: Pedersen, Louise Henriette
; APPLICANT: Rasmussen, Peter Birk
; TITLE OF INVENTION: Novel Application of Vaccination Against TNF-alpha
; FILE REFERENCE: 674542-2020
; CURRENT APPLICATION NUMBER: US/10/939,107
; CURRENT FILING DATE: 2004-09-10
; PRIOR APPLICATION NUMBER: PCT/DK03/00147
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: 60/363,128
; PRIOR FILING DATE: 2002-03-11
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 2773
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p2Op2F expression vector for insect cells
; NAME/KEY: misc_recomb
; LOCATION: (561)..(566)
; OTHER INFORMATION: HindIII site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (573)..(578)
; OTHER INFORMATION: Aval site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (586)..(591)
; OTHER INFORMATION: EcoRI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (593)..(598)
; OTHER INFORMATION: BamHI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (625)..(630)

Query Match	6.0%	Score 33.8	DB 11	Length 600
Best Local Similarity	54.4%	Pred. No. 1.5		
Matches 68	Conservative	0	Mismatches 57	Indels 0
Gaps				0

Query Match
6.0%; Score 33.8; DB 2; Length 3057;

Best Local Similarity 54.4%; Pred. No. 2.6;
Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 10 AACATGATGGTCTAATGTTCTTCAACAACAATTCGTTGAACGTGTTTTCATGT 69
Db 2926 AATAATGTTGGTGAATGTTTCATTAACATCATTAACAAATTTAATGAAAAAGCATGC 2867

QY 70 TTGCCAACAGCACCTTTTATCTCGGTGGCCTCCCAACCACTTTTTCACCTGCAA 129
Db 2866 ATTGTACAGATAGATACCCAGGTATATTAGTCCCATGATTTTTTTGAATGTAT 2807

QY 130 AAAAA 134
Db 2806 CTAAG 2802

RESULT 8
US-60-655-875-26427
; Sequence 26427, Application US/60655875
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/60/655,875
; CURRENT FILING DATE: 2005-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 26427
; LENGTH: 1391
; TYPE: DNA
; ORGANISM: Heterodera glycines
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1051)..(1051)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1380)..(1381)
; OTHER INFORMATION: n is a, c, g, or t
US-60-655-875-26427

Query Match
Best Local Similarity 6.0%; Score 33.6; DB 13; Length 1391;
Matches 48; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 98 GCCTCCCAACCAACTTTTTCACCTGCAAAAAACACGCTTTTGACGCGGCCCAT 157
Db 426 GCCCCCCCAACCAACAATTTTTTACAAAAAATTTTTTCTCCCGGGGCAC 485

QY 158 ACATAGTACAAA 169
Db 486 AAAAAACAAAA 497

RESULT 9
US-10-472-963-1098
; Sequence 1098, Application US/10472963
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS954PCT
; CURRENT APPLICATION NUMBER: US/10/472,963
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: PCT/US02/09370

; PRIOR FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 09/950,082
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US 09/950,083
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 1834
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1098
; LENGTH: 43053
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-472-963-1098

Query Match
Best Local Similarity 56.4%; Pred. No. 9.1;
Matches 62; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 37 CAACAACAATTCGTTGAACGTGTTTTCATGTTTGCCAAACAGCACCTTTTACTCGGT 96
Db 22475 CAACAACCTCTCTGTTCCATTGGGACTTTCAGTTGCGGAACCTCACCTTTTACAGTCTGT 22534

QY 97 GGCCTCCCAACCAACTTTTTCACCTGCAAAAAACACGCTTTTGCA 146
Db 22535 CATCTTCCAGCCTTAACCTCAGTGGTCCACAATTAAGCAAAATTTTGCA 22584

RESULT 10
US-10-472-963-1097
; Sequence 1097, Application US/10472963
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS954PCT
; CURRENT APPLICATION NUMBER: US/10/472,963
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: PCT/US02/09370
; PRIOR FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 09/950,082
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US 09/950,083
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 1834
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1097
; LENGTH: 43056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-472-963-1097

Query Match
Best Local Similarity 56.4%; Pred. No. 9.1;
Matches 62; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 37 CAACAACAATTCGTTGAACGTGTTTTCATGTTTGCCAAACAGCACCTTTTACTCGGT 96
Db 22480 CAACAACCTCTCTGTTCCATTGGGACTTTCAGTTGCGGAACCTCACCTTTTACAGTCTGT 22539

QY 97 GGCCTCCCAACCAACTTTTTCACCTGCAAAAAACACGCTTTTGCA 146
Db 22540 CATCTTCCAGCCTTAACCTCAGTGGTCCACAATTAAGCAAAATTTTGCA 22589

RESULT 11
US-11-079-463-3281
; Sequence 3281, Application US/11079463
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FRA
; FILE REFERENCE: PS954PCT
; CURRENT APPLICATION NUMBER: PCT/US02/09370

```
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 3281
; LENGTH: 1221
; TYPE: DNA
; ORGANISM: B.fragilis
US-11-079-463-3281

Query Match          5.7%; Score 32.4; DB 11; Length 1221;
Best Local Similarity 55.3%; Pred. No. 5.3;
Matches 63; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 29 TGTGCTTCAACAACAAATCTGTGAACTGTTTCATGTTTGCACCAAGCACCTTTA 88
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1085 TGTGTTTCCACAATGTCGTGCGATCTCTCGGCAACTTTGCGCATAGGTATATTT 1144

QY 89 TACTCGGTGCGCTCCCAACCACTTTTTCACCTGCAAAAAACACGCTTT 142
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1145 TATCTATTGGTGTCTTCCATCAACGTATTATATCTGAAAAAGAGATGTTGT 1198

RESULT 12
US-11-079-045-3281
; Sequence 3281, Application US/11079045
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; FILE REFERENCE: PATH00-03DIV1
; CURRENT APPLICATION NUMBER: US/11/079,045
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 3281
; LENGTH: 1221
; TYPE: DNA
; ORGANISM: B.fragilis
US-11-079-045-3281

Query Match          5.7%; Score 32.4; DB 11; Length 1221;
Best Local Similarity 55.3%; Pred. No. 5.3;
Matches 63; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 29 TGTGCTTCAACAACAAATCTGTGAACTGTTTTCATGTTTGCACCAAGCACCTTTA 88
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1085 TGTGTTTCCACAATGTCGTGCGATCTCTCGGCAACTTTGCGCATAGGTATATTT 1144

QY 89 TACTCGGTGCGCTCCCAACCACTTTTTCACCTGCAAAAAACACGCTTT 142
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1145 TATCTATTGGTGTCTTCCATCAACGTATTATATCTGAAAAAGAGATGTTGT 1198

RESULT 13
US-11-079-143-25442
; Sequence 25442, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
```

```
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25442
; LENGTH: 1268
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-25442

Query Match          5.7%; Score 32.4; DB 10; Length 1268;
Best Local Similarity 51.4%; Pred. No. 5.4;
Matches 75; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 327 ACATTATCGGACGGACGAGTGTTCCTTATCGTGCAGGACGCGCTCCATATCAGCCGCGG 386
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 617 ACTGTACATAACTGTGGAGGCTGGGTGTCCTATGTGGAGCCACAACAGCTTCCGGTGGAGG 676

QY 387 TAACCGCAGCCGCGACGAACTCCTTATCGGAACAGGACGCGCTCCATATCAGCCGCGG 446
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 677 TGCCACACAGCGGCTTACTCTTGGCGGCTCCGCTGGGTTGGACATTCGCCATACA 736

QY 447 TTATCTCATGCGGTGACCGGACAGC 472
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 737 AATCTCTCAAGTCCAGAAAGTACAAG 762

RESULT 14
US-11-097-143-38495
; Sequence 38495, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38495
; LENGTH: 1342
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 03:45:00 ; Search time 3951.56 Seconds
(without alignments)
5432.847 Million cell updates/sec

Title: US-09-896-888A-1
Perfect score: 564
Sequence: 1 catgatgataacaatgtat.....tggtagcgacacaacatg 564

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81.2	14.4	679	1 AV398660	AV398660 AV398660
2	39.6	7.0	954	9 CL077394	CL077394 CH216-145
3	39	6.9	762	9 BX147578	BX147578 Danio rer
4	39	6.9	1324	9 CG751412	CG751412 P045-4-E0
5	38.6	6.8	758	8 BH517933	BH517933 BGGBL87TF
6	38.2	6.8	535	7 N63392	N63392 yz35h10.s1
7	38	6.7	578	8 BZ902838	BZ902838 CH240.23B
8	38	6.7	582	8 BZ899636	BZ899636 CH240.16P
9	37.2	6.6	425	6 CB410241	CB410241 NISC.nc10
10	37.2	6.6	669	7 CO384338	CO384338 AGENCOURT
11	36.8	6.5	1101	9 CNS0039X	AL063938 Drosophil
12	36.6	6.5	432	1 AI434744	AI434744 fb34506.x
13	36.6	6.5	580	4 B1843287	B1843287 ft59f05.x
14	36.6	6.5	1048	3 CR709617	CR709617 Tetraodon
15	36.4	6.5	480	6 CB484581	CB484581 cclwtbho
16	36.4	6.5	857	9 CG935894	CG935894 MBEAV05TR
17	35.8	6.3	461	8 BH757407	BH757407 SALK_0561
18	35.8	6.3	655	9 AG358704	AG358704 Mus muscu
19	35.8	6.3	675	4 BG695019	BG695019 NISC.i.v11
20	35.8	6.3	715	2 BE374425	BE374425 601237596
21	35.8	6.3	1049	9 CL510716	CL510716 SAIL_836
22	35.6	6.3	516	4 BM278655	BM278655 AS_tgr_65
23	35.6	6.3	613	9 CL753134	CL753134 OR_BBa012
24	35.6	6.3	626	9 AG401197	AG401197 Mus muscu

25	35.6	6.3	665	9 CL742215	CL742215 OR_BBa007
26	35.6	6.3	787	8 AQ857977	AQ857977 nbeB0011D
C 27	35.6	6.3	1080	9 CNS02AEV	AL188860 Tetraodon
C 28	35.4	6.3	755	8 AQ751255	AQ751255 HS_5574_B
29	35.4	6.3	835	7 CK706618	CK706618 ZF101-P00
30	35.4	6.3	997	9 CG020891	CG020891 ZMMBRC-055
31	35.2	6.2	815	9 CC861440	CC861440 NDL.122K2
C 32	35.2	6.2	1019	4 BG247625	BG247625 602359311
C 33	35.2	6.2	1101	9 CNS00LOO	AL068607 Drosophil
C 34	35	6.2	589	4 BMS25801	BM525801 sak72a09.
C 35	35	6.2	780	9 AG394612	AG394612 Mus muscu
C 36	35	6.2	988	9 CNS0067B	AL062985 Drosophil
C 37	34.8	6.2	771	7 CN755468	CN755468 IDOAAAL5D
C 38	34.8	6.2	851	7 CO808859	CO808859 AGENCOURT
C 39	34.8	6.2	916	8 BH150387	BH150387 ENTPMO1TF
C 40	34.8	6.2	1275	9 CL645025	CL645025 CH213-86H
41	34.6	6.1	717	9 CE275017	CE275017 tigr-gss-
C 42	34.6	6.1	892	9 CNS03TCN	AL259808 Tetraodon
C 43	34.6	6.1	1123	9 CL089685	CL089685 ISB1-13L5
C 44	34.4	6.1	249	9 CG578872	CG578872 OST216624
45	34.4	6.1	455	2 BE332182	BE332182 uc09a08.y

ALIGNMENTS

RESULT 1
LOCUS AV398660 679 bp mRNA linear EST 05-FEB-2000
DEFINITION AV398660 Bombyx mori ovary BmNPV infected; 6 hr after inoculation
Bombyx mori cDNA clone NV060140 T3, mRNA sequence.
ACCESSION AV398660
VERSION AV398660.1 GI:6902312
KEYWORDS EST.
SOURCE Bombyx mori (domestic silkworm)
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea; Bombycidae; Bombyx.
REFERENCE 1 (bases 1 to 679)
Mita,K., Moriyono,M., Shimada,T., Okano,K. and Maeda,S.
Bombyx mori cDNA
Unpublished (2000)
CONTACT: Mita K
Genome Research Group
National Institute of Agrobiological Sciences
Owashi 1-2, Tsukuba, Ibaraki 305-8634, Japan
Email: kmica@nias.affrc.go.jp
Method: uni-directional, sequence direction: sequenced from T3 primer (5' -> 3')

FEATURES
source
/organism="Bombyx mori"
/mol_type="mRNA"
/db_xref="taxon:7091"
/clone="NV060140"
/tissue_type="ovary"
/cell_type="BmN cultured cell"
/dev_stage="BmNPV infected; 6 hr after inoculation"
/clone_lib="Bombyx mori ovary BmNPV infected; 6 hr after inoculation"

ORIGIN
Query Match 14.4%; Score 81.2; DB 1; Length 679;
Best Local Similarity 71.3%; Pred. No. 2.1e-14;
Matches 107; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 129 AAAAAACGCTTTTCACGCGGGCCATACATAGTACAACTCTAGCTTCGTAGCTA 188
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 AAAAAAGTTCGCTATGTTTGACATATATATACAGTACGAACCTCTACAAATCTAGCTA 60

```

QY 189 TTTTACATAAATAGTCTACACGGTTGTATACGCTCCAAATACACTACACACATTGAACC 248
Db 61 TTTTATTAGATGCTACACTGTACTATACCGCTCTCAATATACTACTACTATCAACT 120

QY 249 TTTTTCGCACTGCAAAAAAGTACGCTGTCGCG 278
Db 121 TTTTTCGATTACAAAAAGTTCAATTTTGC 150

RESULT 2
LOCUS CL077394
DEFINITION CH216-145B11_Sp6.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CH216-145B11, genomic survey sequence.
VERSION CL077394
KEYWORDS GSS.
SOURCE CL077394.1 GI:40533307
ORGANISM Xenopus tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 954)
AUTHORS Krenitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp6 ATTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 163
High quality sequence stop: 226.
FEATURES
Location/Qualifiers
1..954
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strains="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-145B11"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/notes="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"
ORIGIN
Query Match 7.0%; Score 39.6; DB 9; Length 954;
Best Local Similarity 52.4%; Pred. No. 0.48;
Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 103 CCCACACCAACTTTTTCGCTGCAAGCAAAACAGCTTTTGCAGCGGGCCCATACATA 162
Db 330 CGCCCCCCCCCTTTTTCACAAAAAACAACCCCTTCTCTCCACCTTAAACATA 389

QY 163 GTACAACTCTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGCT 222
Db 390 ATCTGAACCTCTATTATCACTCATCTATTATATACGGATAATGTACCCCTACTGTAAATC 449

QY 223 CCAATACACTACACACATTGAACCTTTTTCAGTGCAGCAAAAGT 268
Db 450 ATACGATATTACCACTCACTGACCTCTCTGTACCCATATAAATT 495

RESULT 3
BX147578/C
LOCUS BX147578
DEFINITION Danio rerio genomic clone DKEY-109J14, genomic survey sequence.
ACCESSION BX147578

```

```

VERSION BX147578.1 GI:27978953
KEYWORDS GSS.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 762)
AUTHORS Humphray, S.J., Huckle, E. and Durham, J.L.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquere@sanger.ac.uk Unpublished
COMMENT This sequence was generated from the T7 end of BAC 109J14. 109J14
is part of the Daniokey BAC library created by R. Plasterk and N.V.
Keygene. Further details:
http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES
Location/Qualifiers
1..762
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-109J14"
/tissue type="Testis"
/notes="Vector pIndigoBAC-536"
ORIGIN
Query Match 6.9%; Score 39; DB 9; Length 762;
Best Local Similarity 56.7%; Pred. No. 0.71;
Matches 72; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 157 TACATAGTACAACTCTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTA 216
Db 655 TATAAATACACAAAAATCCTATTAGTCGAGCAATTTATTTAAATCGTTTATAGCAAAAC 596

QY 217 TACGCTCCAAATACACTACACACATGAACTTTTTCAGTGCAGCAAAAGTACGTGTGC 276
Db 595 TCATTTCCAAAAACACTACTAGTCAATGTCAGTCCAGGAATGCAATCGTTAATATGCA 536

QY 277 GCAGTCA 283
Db 535 GTCGACA 529

RESULT 4
CG751412 1324 bp DNA linear GSS 24-OCT-2003
LOCUS P045-4-E06.ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,
DEFINITION genomic survey sequence.
ACCESSION CG751412.1 GI:37973841
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 1324)
AUTHORS Srinivasan, J., Sinz, W., Jesse, T., Wiggers-Perebolte, L., Jansen, K.,
Buntjer, J., van der Meulen, M. and Sommer, R.J.
TITLE An integrated physical and genetic map of the nematode Pristionchus
JOURNAL pacificus
MEDLINE Mol. Genet. Genomics 269 (5), 715-722 (2003)
PUBMED 12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: raif.sommer@uebingen.mpg.de
Class: BAC ends.
FEATURES
Location/Qualifiers

```

```

source
1. .1324
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/notes="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN
Query Match 6.9%; Score 39; DB 9; Length 1324;
Best Local Similarity 48.1%; Pred. No. 0.81;
Matches 111; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 38 AACACAAATCTGTGAACCTGTTTTCATGTTTGCACCAAGCAGCCTTTATACCTGGTG 97
DB 1092 AATATACACACTTTTAACTTTTTTTTTTTTTTTTCTGAACCCCTTTTAAACCTCC 1151

QY 98 GCCTCCCAACCAACTTTTTTGCACTGCAAAAAACAGCTTTTGCAGCGGGCCCAT 157
DB 1152 TCCTTAGTAAACCCCAAGTTTGAATTCATATACGAGCCCTTTAAATAATTAACCTCA 1211

QY 158 ACATAGTACAACTCTACGTTTCGTAGACTATTTTACATTAATAGTCTACACCGTTGTAT 217
DB 1212 AGAAATTCACACTAGCTAAAGGGAATTTATTTCAAAAAAATAATATTCCTTTATT 1271

QY 218 ACCTCCAAATACACTACACACATTCGAACCTTTTTCAGTCGCAAAAAAGT 268
DB 1272 TGGAAAAAACAATTTTCCACCCCAATATTTTTTTTGGCCCAAAAAAGT 1322

RESULT 5
BH517933/c
LOCUS BH517933 758 bp DNA linear GSS 13-DEC-2001
DEFINITION BOGBL877P BOGB Brassica oleracea genomic clone BOGBL87, genomic
survey sequence.
ACCESSION BH517933
VERSION BH517933.1 GI:17726023
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 758)
AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TP
Class: sheared ends.
FEATURES
Location/Qualifiers
1. .758
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TOL000DH3"
/db_xref="taxon:3712"
/clone_lib="BOGBL87"
/notes="Vector: pBOS1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pBOS1 using BstXI linkers"

ORIGIN
Query Match 6.8%; Score 38.6; DB 8; Length 758;
Best Local Similarity 52.1%; Pred. No. 0.95;
Matches 86; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 110 CCAACTTTTTTGCACGCAAAAAACACGCTTTTGCACGCGGCCCATACATAGTACAA 169
DB 527 CCAGTTGGTTTGCACCTTTCAAGATTAAATGATCAGTACTCTCATCAACAAATATTATAA 468

QY 170 CTCCTAGCTTTCTAGACTATTTTACATAAATAGTCTACACCGTTGTATAGCTCCAATA 229
DB 467 CTCCTCAAAATTAATTAATACGATAAATTAATAAAAAAATAATATGATGATAGGTTAATA 408

QY 230 CACTACCAACATTTGAACCTTTTTCAGTGCAGAAAAAGTACGTGT 274
DB 407 AAATTCATACATTAATTAATATGTAAAAAATAAATAATATAGACATTT 363

RESULT 6
N63392 535 bp mRNA linear EST 01-MAR-1996
LOCUS yz3sh10.s1 Morton Fetal Cochlea Homo sapiens cDNA clone
DEFINITION IMAGE:285091 3', mRNA sequence.
ACCESSION N63392
VERSION N63392.1 GI:1211221
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 535)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.P., Chiapelli,B.,
Chisoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
Trevaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R.
and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
PUBMED 8889549
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estowatson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: ml3 -40 forward
High quality sequence stop: 246.
Location/Qualifiers
1. .535
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3892208"
/db_xref="taxon:9606"
/clone_lib="IMAGE:285091"
/tissue_type="cochlea"
/dev_stage="16-22 week fetus"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_lib="Morton Fetal Cochlea"
/notes="Organ: ear; Vector: pBluescript SK-; Site 1: EcoRI;
Site 2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned
unidirectionally. Primer: Oligo dT. Fetal cochlea, normal.
37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP
XR Vector. Library constructed by N. Robertson, C. Morton.
-5' adaptor sequence: 5' GAATTCGACGACGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

ORIGIN
Query Match 6.8%; Score 38.2; DB 7; Length 535;
Best Local Similarity 49.2%; Pred. No. 1.2;
Matches 97; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 156 ATACATAGTACAAACTCTACGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGT 215

```

```

Db      218 ATGTATACCTCTAACAGTTAGCTTGTATAAAAAAATCAACAAAAAGCAAAACCATTTT 277
Qy      216 ATACGCTCCAAATACACTACACACATGAACTTTTTCAGTGCAGAAAAAGTAGCTGTC 275
Db      278 ATTTGCCCAAATTTTACCAGGTTTTTACTTTAAAAAATTTCTACCTTATGGGTAAATCCAGGC 337
Qy      276 GGCAGTTCAGTAGGCGGCTTATCGGTCGCGTCTGTCAGTACGATACATTATCG 335
Db      338 TCCAATCGGTAAACAACCTTGGCATGTTTTTACCTACCACTGGCACCAGGTTTTAG 397
Qy      336 GACCGGACGAGTGTGT 352
Db      398 GACNGGATAAATNAITTT 414

```

```

RESULT 7
BZ902838/c
LOCUS   BZ902838
DEFINITION CH240_23B3.TJ CHORI-240 Bos taurus genomic clone CH240_23B3,
genomic survey sequence.
ACCESSION BZ902838
VERSION   BZ902838.1 GI:31627927
KEYWORDS  GSS.
SOURCE    Bos taurus (cow)
ORGANISM  Bos taurus

```

```

REFERENCE
AUTHORS  Larkin,D.M., Everts-van der Wind,A., Rebeiz,M., Schweitzer,P.,
Bachman,S., Green,S., Campos,E.J., Benson,L.D., Edwards,J., Liu,L.,
Womack,J.E., de Jong,P.J. and Lewin,H.A.
TITLE    A Cattle-Human Comparative Map Built with Cattle BAC-ends and Human
Genome Sequence
JOURNAL  Unpublished (2003)
COMMENT  Other GSSs: CH240_23B3.TV
Contact: Harris Lewin
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by by University of Illinois at Urbana
Champaign, USA with funds provided by grant No. AG202-34480-11828
from USDA-CSREES and AG99-35205-8534 from USDA/NRI (Livestock
Genome Sequencing Initiative)
Plate: 23 row: B column: 3
Seq primer: SP6
Class: BAC ends.

```

```

FEATURES
source
1..578
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_23B3"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

```

```

ORIGIN
Query Match 6.7%; Score 38; DB 8; Length 578;
Best Local Similarity 55.2%; Pred. No. 1.4;

```

```

Best Local Similarity 55.2%; Pred. No. 1.4;
Matches 74; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
Qy      1 CATGATGATAAACAATGATGTCGTAATGTTCTTCAACAACAATTCGTTGAACGTGTG 60
Db      252 CATGCTGATCAATGAGGCATAAAGATATAATTTGGTTTATAGACACATTTGTTGGAATGTA 193
Qy      61 TTTTCATGTTTTGCCCAACAAGCACCTTTATATCTCGTGGCCTCCCAACCAACTTTTTT 120
Db      192 TTTTACCAGCTTTTAAACACCAACATTTTATGTTTGGGCTTAAATCATCAGCAAAATGAAT 133
Qy      121 GCACTGCAAAAAAAA 134
Db      132 TCAGCATTAATAAAA 119

```

```

RESULT 8
BZ899636
LOCUS   BZ899636
DEFINITION CH240_16P5.TV CHORI-240 Bos taurus genomic clone CH240_16P5,
genomic survey sequence.
ACCESSION BZ899636
VERSION   BZ899636.1 GI:31624687
KEYWORDS  GSS.
SOURCE    Bos taurus (cow)
ORGANISM  Bos taurus

```

```

REFERENCE
AUTHORS  Larkin,D.M., Everts-van der Wind,A., Rebeiz,M., Schweitzer,P.,
Bachman,S., Green,S., Campos,E.J., Benson,L.D., Edwards,J., Liu,L.,
Womack,J.E., de Jong,P.J. and Lewin,H.A.
TITLE    A Cattle-Human Comparative Map Built with Cattle BAC-ends and Human
Genome Sequence
JOURNAL  Unpublished (2003)
COMMENT  Contact: Harris Lewin
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by by University of Illinois at Urbana
Champaign, USA with funds provided by grant No. AG202-34480-11828
from USDA-CSREES and AG99-35205-8534 from USDA/NRI (Livestock
Genome Sequencing Initiative)
Plate: 16 row: P column: 5
Seq primer: T7
Class: BAC ends.

```

```

FEATURES
source
1..582
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_16P5"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

```

```

ORIGIN
Query Match 6.7%; Score 38; DB 8; Length 582;
Best Local Similarity 55.2%; Pred. No. 1.4;

```

Matches	74; Conservative	0; Mismatches	60; Indels	0; Gaps	0;
Qy	1	CATGATGATAAACAATGATGGTAAATGTTGCTTCAACAACAATTCGTGTGAACGTG	60		
Db	265	CATGCTGATCAATGAGGCATAAAGATATATATGTTTATAGACACATTTGTTGGAATGTA	324		
Qy	61	TTTTTCATGTTTGCACACAGCACCTTTATACCTGGTGGCTCCACCACCAACTTTTTT	120		
Db	325	TTTACCAAGCTTTTAAACACCACTTATGTTTGGGCGCTAAATCATCAGCAATGAATT	384		
Qy	121	GCACTGCAAAAAA	134		
Db	385	TCAGCATTAARAAA	398		
RESULT 9					
LOCUS	CB410241	425 bp	mRNA	linear	EST 24-MAR-2003
DEFINITION	NISC_nc10d09.x1 COGENE 6E MAX Homo sapiens cDNA clone IMAGE:5776553 3', mRNA sequence.				
ACCESSION	CB410241				
VERSION	CB410241.1	GI:29166981			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 425)				
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.				
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov cDNA Library Preparation: CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC) Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov Plate: LLAM12849 row: G column: 18 Seq primer: -21M13 forward primer (ABI).				
FEATURES	Location/Qualifiers				
source	1..425				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:5776553"				
	/tissue_type="maxilla, pooled"				
	/dev_stage="6 weeks postconception"				
	/lab_host="DH10B"				
	/clone_lib="COGENE 6E MAX"				
	/note="Vector: pAMP1; cDNA primed using oligo-dT primer, directionally cloned into UDG sites of pAMP1. Size selected for insert sizes ranging from 0.2-1.8 kb. Normalized to Cot5. Primary library, non-amplified. Library constructed by M. Lovett. For more information on this library, please contact R. Tidwell (Washington University) or visit the COGENE website at http://hg.wustl.edu/COGENE/."				
ORIGIN					
Query Match	6.6%; Score 37.2; DB 6; Length 425;				
Best Local Similarity	54.3%; Pred. No. 2.3;				
Matches	75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;				
Qy	115	TTTTTTGCACTGCAAAAAACAGCTTTTTCAGCGCGGCCCATACATAGTACAACTCTA	174		
Db	66	TTTTTTCCAAAGGAAAAAATTTTTCCTCGGGTTTATCCACTGTCATACTGTA	125		
Qy	175	CGTTTCGTAGACTATTTTACATAAATAGTCTACACCGTTGTATACGTCCTCAATACACTA	234		

Db	126	TTTTTCATGCAATATATTTGCCAAAAAACTCAGCTTTTATTTTCCATTTTAAACACTA	185		
Qy	235	CCACACATTTGAACCTTTT	252		
Db	186	CAATATTTACAAGCTGTT	203		
RESULT 10					
LOCUS	CO384338	669 bp	mRNA	linear	EST 30-JUN-2004
DEFINITION	AGENCOURT 26189343 Blumberg Cho dorsal blastopore lip Xenopus laevis cDNA clone IMAGE:7297357 3', mRNA sequence.				
ACCESSION	CO384338				
VERSION	CO384338.1	GI:49490161			
KEYWORDS	EST.				
SOURCE	Xenopus laevis (African clawed frog)				
ORGANISM	Xenopus laevis				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus; Xenopus.				
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.				
TITLE	1 (bases 1 to 669)				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)				
COMMENT	Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgapbs-r@mail.nih.gov Tissue Procurement: Bruce Blumberg CDNA Library Preparation: B. Blumberg CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov				
Plate:	LLAM15322 row: j column: 11				
High quality sequence stop:	534.				
FEATURES	Location/Qualifiers				
source	1..669				
	/organism="Xenopus laevis"				
	/mol_type="mRNA"				
	/db_xref="taxon:8355"				
	/clone="IMAGE:7297357"				
	/sex="both"				
	/tissue_type="dorsal blastopore lip"				
	/lab_host="TOP10"				
	/clone_lib="Blumberg Cho dorsal blastopore lip"				
	/note="Organ: embryo; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Library was prepared from 50 ug of total RNA by oligo-dT priming and AMV reverse transcriptase. After addition of EcoRI linkers and EcoRI-XhoI digestion, the cDNA was size selected by chromatography on Sepharose CL-4B columns and fractions containing cDNAs larger than 500 bp were ligated into EcoRI-XhoI-digested lambda ZAPII (UnizAP-XR) and packaged in vitro. Average insert size is 1.4 kb. The original library contained 6 x 106 recombinants, of which 3 x 106 were amplified and stored at -70 C in SM buffer containing 7% DMSO. 3 x 106 pfu were mass excised and the resulting phagemids used to infect Top10F. References: Science 253, 196-196 and Methods in Molecular Biology 97, 555-574. Additional sequences from this library have been deposited under the name Xenopus laevis dorsal blastopore lip. Library constructed by Bruce Blumberg (University of California, Irvine, Department of Developmental and Cell Biology)."				
ORIGIN					
Query Match	6.6%; Score 37.2; DB 7; Length 669;				
Best Local Similarity	57.9%; Pred. No. 2.6;				
Matches	66; Conservative 0; Mismatches 48; Indels 0; Gaps 0;				

153	Qy	CCCATACATAGTACAAACTCTACGTTTCGTAGACATATTTTACATAAATAGTCTACCGT	212
104	Db	CCAAATAATAGGGAAACCTTCGTGTAGTGTGTAAGAATACATCCAAATGTATGCAACGT	163
213	Qy	TGTATACGCTCCAAATACACTACACACATTTGAACCTTTTTGCAAGTCGCAAAAA	266
164	Db	TGTAGTCTCAGAAATACATTTTAATCTTTTTTTCTTTTTGCAAGTTGCAAAATA	217

RESULT 11	
CNS0039X/c	
LOCUS	CNS0039X linear GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC # BACR08012 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL063938
VERSION	AL063938.1 GI:4941795
KEYWORDS	GSS.
SOURCE	Drosophila melanogaster (fruit fly)
ORGANISM	Drosophila melanogaster
	Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 1101)
AUTHORS	Genoscope.

TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammos in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial, EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and PST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <http://bacpac.med.buffalo.edu/drosophila.bac.htm>.

FEATURES	source
1. 1101	
Location/Qualifiers	
/organism="Drosophila melanogaster"	
/mol_type="genomic DNA"	
/db_xref="taxon:7227"	
/clone="BACR08012"	
/clone_lib="RPC1-98"	
/note="end : T7"	

Query Match	6.5%	Score 36.8;	DB 9;	Length 1101;
Best Local Similarity	35.9%;	Pred. No. 3.9;		
Matches 61;	Conservative	35;	Mismatches	74;
			Indels	0;
			Gaps	0;

[illegible]

QY 133 AACACGCTTTTTCACGCGGGCCCATACATAGTACAAACTCTACGTTTCGTAGACTATTTT 192

Db 1008 CMCACYYYWCAACAACMCMTCCCTCTCTATAATAACATCTAATAATCAWMAAYTACMAH 949

Qy	193	ACATAAATAGTCTACACCGTTGTATTACGCTCAAAATACACTACCACAT	242
Db	948	MAMAAWAWATAWCWMAHYHMTTCACACACWCACCTATCATCWMACCCYACWT	899

RESULT 12
AI437474

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Research	1980, Vol. 83, No. 2, pp. 11-20
3. The Importance of Parental Involvement	Journal of Educational Research	1980, Vol. 83, No. 3, pp. 21-30
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 31-40
5. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 5, pp. 41-50
6. The Impact of Teacher Education on Classroom Practice	Journal of Educational Research	1980, Vol. 83, No. 6, pp. 51-60
7. The Importance of Teacher Collaboration	Journal of Educational Research	1980, Vol. 83, No. 7, pp. 61-70
8. The Effect of Teacher Attitudes on Student Behavior	Journal of Educational Research	1980, Vol. 83, No. 8, pp. 71-80
9. The Role of the School in the Development of the Child	Journal of Educational Research	1980, Vol. 83, No. 9, pp. 81-90
10. The Impact of Teacher Training on Student Learning	Journal of Educational Research	1980, Vol. 83, No. 10, pp. 91-100

FEATURES source

ORIGIN

Query Match 6.5%; Score 36.6; DB 1; Length 432;
Best Local Similarity 50.9%; Pred. No. 3.7;
Matches 87; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 109 ACCAACTTTTTTGCATCGAAAAAACAGCTTTTGCACGCGGGGCCCATACATAGTACAA 168

Db 21 ATCACTTATTTCTACTCGGTACAGCGTGAATTTTACATTTTTCGAAAGTGACAAACTG 80
 Qy 169 ACTCTAGCTTTCTAGTACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAAT 228
 Db 81 AGTGACAAATTCACAAATATATAAACAACACTGTTCTGTGGCTACTAAGCACTT 140
 Qy 229 AACTACCAACATGTAACCTTTTTCAGTGCAAAAAGTACGTGCGCA 279
 Db 141 ACTCTACTACACACTTCAATTCACGAGAGCGCTAAAGTACATGACAGAA 191

RESULT 13

BI843287
 LOCUS ft59f05.x1 Gong zebrafish ovary Danio rerio cDNA clone
 DEFINITION IMAGE:5157656 3', mRNA sequence.

ACCESSION BI843287.1 GI:15955810

VERSION EST

SOURCE Danio rerio (zebrafish)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE

1 (bases 1 to 580)
 Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Streptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

WashU Zebrafish EST Project 1998

Unpublished (1998)

TITLE

Contact: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrfish@watson.wustl.edu

The library was constructed by Dr. Z. Gong. DNA Sequencing by: Washington University Genome Sequencing Center St. Louis. Please contact Zhiyuan Gong for further information on this library (National University of Singapore, Department of Biological Sciences, Lower Kent Ridge Road, Singapore 119260).

Seq primer: T7 from Gibco

High quality sequence stop: 410.

FEATURES

source

1. 580
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:5157656"
 /sex="Female"
 /dev_stage="4-5 month"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Gong zebrafish ovary"

/note="Organ: ovary (pooled); Vector: pBluescript SK-; Site1: XhoI; Site2: EcoRI; Poly A+ RNA was isolated from the ovaries of 2 female adult zebrafish (4-5 month old). cDNAs were made using oligo-dT primers and inserted into lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo mass-excision to pBluescript SK- following the Washington University protocol (http://genome.wustl.edu/est/lambda_protocol.shtml). Please contact Zhiyuan Gong for further information on this library (National University of Singapore, Department of Biological Sciences, Lower Kent Ridge Road, Singapore 119260)."

ORIGIN

Query Match 6.5%; Score 36.6; DB 4; Length 580;
 Best Local Similarity 50.9%; Pred. No. 3.9;

Matches 87; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
 Qy 109 ACCAACTTTTTCGACTGCAAAAAACACGCTTTTTCACGCGGCCCATATAGTACAA 168
 Db 27 ATCACTTTTATTTCTACTGTGTACACGCGTGAATTTTACTTTTGCAGAGTGACAAACTG 86
 Qy 169 ACTCTAGCTTTCTAGTACTATTTTACATAAATAGTCTACACCGTTGTATACGCTCCAAAT 228
 Db 87 AGTGACAAATTTTACAAAAATATATAAACAACACTGTTTCTGTGGCTACTAAGCACTT 146
 Qy 229 AACTACCAACATGTAACCTTTTTCAGTGCAAAAAGTACGTGCGCA 279
 Db 147 ACTCTACTACACACTTCAATTCACGAGAGCGCTAAAGTACATGACAGAA 197

RESULT 14

CR709617/c
 LOCUS CR709617 1048 bp mRNA linear HTC 19-AUG-2004
 DEFINITION Tetraodon nigroviridis full-length cDNA.

ACCESSION CR709617

VERSION CR709617.1 GI:51207526

KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.

SOURCE Tetraodon nigroviridis

ORGANISM Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.

1 (bases 1 to 1048)

Genoscope.

Direct Submission

TITLE Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage

JOURNAL : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.

More information available at

http://www.genoscope.cns.fr/tetraodon.

FEATURES

source

1. 1048
 /Location/Qualifiers
 /organism="Tetraodon nigroviridis"
 /mol_type="mRNA"
 /db_xref="taxon:9883"
 /tissue_type="Eggs"

ORIGIN

Query Match 6.5%; Score 36.6; DB 3; Length 1048;
 Best Local Similarity 49.5%; Pred. No. 4.5;
 Matches 93; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
 Qy 22 GTGCTAATGTGTCTTCAACAACAATTTCTGTGAACTGTGTTTCATGTGTTGCCAACAGC 81
 Db 1033 GTGGAAGATTGTTTATTTTACACTGCCATTCACATGTGCTCTGTGGTTTGGACAGACTC 974
 Qy 82 ACCTTTATCTCGTGGCGCTCCCAACCAACCACTTTTTCACCTGCAAAAACACGCTT 141
 Db 973 GACCATCGACTGGGCTCCAAATCATCCACTAGCGCAAGAAAATCCATTTTATTTACAGCA 914
 Qy 142 TTGCACGCGGGCCCATACATAGTACAACTCTACGTTTCGTAGACTATTTTACATAATA 201
 Db 913 GTTAATGCGACTAACCACTTAATCAAAATGAGTTTATTAGTTTTTTTATTATAA 854
 Qy 202 GTCTACAC 209
 Db 853 CTCTTAC 846

RESULT 15

CB484581
 LOCUS cclwtbh009074 Coregonus clupeaformis head Coregonus clupeaformis
 DEFINITION cDNA, mRNA sequence.
 ACCESSION CB484581
 VERSION CB484581.1 GI:29295807

ORIGIN

Query Match 6.5%; Score 36.6; DB 4; Length 580;
 Best Local Similarity 50.9%; Pred. No. 3.9;

KEYWORDS EST.
SOURCE Coregonus clupeaformis
ORGANISM Coregonus clupeaformis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Coregonus.
REFERENCE 1 (bases 1 to 480)
AUTHORS GRASP Consortium, Davidson, W.S., Koop, B.P. and
http://web.uvic.ca/cbr/grasp.
TITLE A survey of Salmo salar transcripts from high complexity cDNA
libraries
JOURNAL Unpublished (2002)
COMMENT Contact: Koop BF
Centre for Biomedical Research
University of Victoria
PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
Tel: 250 472 4067
Fax: 250 472 4075
Email: bkoop@uvic.ca
Centre for Biomedical Research, University of Victoria cDNA
preparation and sequencing: Roberto Alberto, Marianne
Beetz-Sargent, Maura Busby, Peter Hunt, Linda McKinnel, BF Koop.
Bioinformatics: Gordon D Brown
POLYA=Yes.

FEATURES Location/Qualifiers
source 1..480
/organism="Coregonus clupeaformis"
/mol_type="mRNA"
/db_xref="taxon:59861"
/clone_lib="Coregonus clupeaformis head"
/note="Vector: pBlueScriptIIISK+; Library Creator:
Kristian R von Schalburg ; Lake whitefish tissue
contributor: Louis Bernatchez (Universite Laval)"

ORIGIN

Query Match 6.5%; Score 36.4; DB 6; Length 480;
Best Local Similarity 51.9%; Pred. NO. 4.4;
Matches 82; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 49 TGTTCAGCTGTTTTCATGTTTGGCCACAGCACCTTTTACTCGGTGGCTCCACC 108
Db 10 TTTTGAATGTCCTTTATTTGTTCCATCATCCACGTTTCTACATCAATCCACATCCA 69
QY 109 ACCAACTTTTTCGCTGCAAAAAACACGCTTTTGCACGGGGCCCATACATAGTACAA 168
Db 70 GACAAATTTATGTTAATATTAACTCACAATTCAGCACTGTATGCTTACTAATACAA 129
QY 169 ACTCTACGTTTCGTAGACTATTTTACATAAATAGTCTA 206
Db 130 CACTTGGGATATGCAAGCCATTTTACAGAAATGGACTA 167

Search completed: May 10, 2005, 07:25:08
Job time : 3957.56 secs

K(00-04-00)